(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 April 2002 (04.04.2002)

PCT

(10) International Publication Number
WO 02/26982 A2

(51) International Patent Classification[7]: C12N 15/12, C12Q 1/68, C07K 14/47, 16/18, A01K 67/027, A61K 38/17, 39/395, B01J 19/00, G01N 33/50, 33/53

(21) International Application Number: PCT/US01/30042

(22) International Filing Date:
25 September 2001 (25.09.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/236,869	29 September 2000 (29.09.2000)	US
60/239,812	11 October 2000 (11.10.2000)	US
60/240,108	12 October 2000 (12.10.2000)	US
60/241,282	17 October 2000 (17.10.2000)	US
60/242,218	20 October 2000 (20.10.2000)	US

(71) Applicant *(for all designated States except US)*: **INCYTE GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94087 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **NGUYEN, Danniel, B.** [US/US]; 1403 Ridgewood Drive, San Jose, CA 95118 (US). **YAO, Monique, G.** [US/US]; 1189 Woodgate Drive, Carmel, IN 40633 (US). **XU, Yuming** [US/US]; 1739 Walnut Drive, Mountain View, CA 94040 (US). **TRIBOULEY, Catherine, M.** [FR/US]; 1121 Tennessee Street #5, San Francisco, CA 94107 (US). **SANJANWALA, Madhu, S.** [US/US]; 210 Sylvia Court, Los Altos, CA 94024 (US). **WALIA, Narinder, K.** [US/US]; 890 Davis Street #205, San Leandro, CA 94577 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **SAPPERSTEIN, Stephanie, K.** [US/US]; 3531 Highland Avenue, Redwood City, CA 94062 (US). **LAL, Preeti** [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). **THORNTON, Michael** [US/US]; 9 Medway Road, Woodside, CA 94062 (US). **GANDHI, Ameena, R.** [US/US]; 837 Roble Avenue #1, Menlo Park, CA 94025 (US). **RAMKUMAR, Jayalaxmi** [US/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). **ELLIOTT, Vicki, S.** [US/US]; 3770 Polten Place Way, San Jose, CA 95121 (US). **ARVIZU, Chandra** [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). **THANGAVELU, Kavitha** [IN/US]; 1950 Montecito Avenue #23, Mountain View, CA 94043 (US). **GIETZEN, Kimberly, J.** [US/US]; 691 Los Huecos Drive, San Jose, CA 95123 (US). **DING, Li** [CH/US]; 3353 Alma Street #146, Palo Alto, CA 94506 (US). **AU-YOUNG, Janice** [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US). **TRAN, Bao** [US/US]; 750 Salberg Avenue, Santa Clara, CA 95051 (US). **POLICKY, Jennifer, L.** [US/US]; 1511 Jarvis Court, San Jose, CA 95118 (US). **LEE, Sally** [US/US]; 825 East Evelyn #425, Sunnyvale, CA 94086 (US). **LU, Dyung, Aina, M.** [US/US]; 233 Coy Drive, San Jose, CA 95123 (US). **BURFORD, Neil** [GB/US]; 105 Wildwood Circle, Durham, CT 06422 (US). **WARREN, Bridget, A.** [US/US]; 10130 Parwood Drive #2, Cupertino, CA 95014 (US). **GURURAJAN, Rajagopal** [IN/US]; 5591 Dent Avenue, San Jose, CA 95118 (US). **DUGGAN, Brendan, M.** [AU/US]; 243 Buena Vista Avenue #306, Sunnyvale, CA 94086 (US). **HONCHELL, Cynthia, D.** [US/US]; 400 Laurel Street #203, San Carlos, CA 94070 (US). **HAFALIA, April, J., A.** [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US).

(74) Agents: **HAMLET-COX, Diana** et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SECRETED PROTEINS

(57) Abstract: The invention provides human secreted proteins (SECP) and polynucleotides which identify and encode SECP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of SECP.

WO 02/26982 A2

SECRETED PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of secreted proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of secreted proteins.

BACKGROUND OF THE INVENTION

Protein transport and secretion are essential for cellular function. Protein transport is mediated by a signal peptide located at the amino terminus of the protein to be transported or secreted. The signal peptide is comprised of about ten to twenty hydrophobic amino acids which target the nascent protein from the ribosome to a particular membrane bound compartment such as the endoplasmic reticulum (ER). Proteins targeted to the ER may either proceed through the secretory pathway or remain in any of the secretory organelles such as the ER, Golgi apparatus, or lysosomes. Proteins that transit through the secretory pathway are either secreted into the extracellular space or retained in the plasma membrane. Proteins that are retained in the plasma membrane contain one or more transmembrane domains, each comprised of about 20 hydrophobic amino acid residues. Secreted proteins are generally synthesized as inactive precursors that are activated by post-translational processing events during transit through the secretory pathway. Such events include glycosylation, proteolysis, and removal of the signal peptide by a signal peptidase. Other events that may occur during protein transport include chaperone-dependent unfolding and folding of the nascent protein and interaction of the protein with a receptor or pore complex. Examples of secreted proteins with amino terminal signal peptides are discussed below and include proteins with important roles in cell-to-cell signaling. Such proteins include transmembrane receptors and cell surface markers, extracellular matrix molecules, cytokines, hormones, growth and differentiation factors, enzymes, neuropeptides, vasomediators, cell surface markers, and antigen recognition molecules. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of The Cell, Garland Publishing, New York, NY, pp. 557-560, 582-592.)

Cell surface markers include cell surface antigens identified on leukocytic cells of the immune system. These antigens have been identified using systematic, monoclonal antibody (mAb)-based "shot gun" techniques. These techniques have resulted in the production of hundreds of mAbs directed against unknown cell surface leukocytic antigens. These antigens have been grouped into "clusters of differentiation" based on common immunocytochemical localization patterns in various

differentiated and undifferentiated leukocytic cell types. Antigens in a given cluster are presumed to identify a single cell surface protein and are assigned a "cluster of differentiation" or "CD" designation. Some of the genes encoding proteins identified by CD antigens have been cloned and verified by standard molecular biology techniques. CD antigens have been characterized as both transmembrane proteins and cell surface proteins anchored to the plasma membrane via covalent attachment to fatty acid-containing glycolipids such as glycosylphosphatidylinositol (GPI). (Reviewed in Barclay, A. N. et al. (1995) The Leucocyte Antigen Facts Book, Academic Press, San Diego, CA, pp. 17-20.)

Matrix proteins (MPs) are transmembrane and extracellular proteins which function in formation, growth, remodeling, and maintenance of tissues and as important mediators and regulators of the inflammatory response. The expression and balance of MPs may be perturbed by biochemical changes that result from congenital, epigenetic, or infectious diseases. In addition, MPs affect leukocyte migration, proliferation, differentiation, and activation in the immune response. MPs are frequently characterized by the presence of one or more domains which may include collagen-like domains, EGF-like domains, immunoglobulin-like domains, and fibronectin-like domains. In addition, MPs may be heavily glycosylated and may contain an Arginine-Glycine-Aspartate (RGD) tripeptide motif which may play a role in adhesive interactions. MPs include extracellular proteins such as fibronectin, collagen, galectin, vitronectin and its proteolytic derivative somatomedin B; and cell adhesion receptors such as cell adhesion molecules (CAMs), cadherins, and integrins. (Reviewed in Ayad, S. et al. (1994) The Extracellular Matrix Facts Book, Academic Press, San Diego, CA, pp. 2-16; Ruoslahti, E. (1997) Kidney Int. 51:1413-1417; Sjaastad, M.D. and Nelson, W.J. (1997) BioEssays 19:47-55.)

Mucins are highly glycosylated glycoproteins that are the major structural component of the mucus gel. The physiological functions of mucins are cytoprotection, mechanical protection, maintenance of viscosity in secretions, and cellular recognition. MUC6 is a human gastric mucin that is also found in gall bladder, pancreas, seminal vesicles, and female reproductive tract (Toribara, N.W. et al. (1997) J. Biol. Chem. 272:16398-16403). The MUC6 gene has been mapped to human chromosome 11 (Toribara, N.W. et al. (1993) J. Biol. Chem. 268:5879-5885). Hemomucin is a novel Drosophila surface mucin that may be involved in the induction of antibacterial effector molecules (Theopold, U. et al. (1996) J. Biol. Chem. 217:12708-12715).

Tuftelins are one of four different enamel matrix proteins that have been identified so far. The other three known enamel matrix proteins are the amelogenins, enamelin and ameloblastin. Assembly of the enamel extracellular matrix from these component proteins is believed to be critical in producing a matrix competent to undergo mineral replacement. (Paine, C.T. et al. (1998) Connect Tissue Res.38:257-267). Tuftelin mRNA has been found to be expressed in human ameloblastoma

tumor, a non-mineralized odontogenic tumor (Deutsch, D. et al. (1998) Connect. Tissue Res. 39:177-184).

Olfactomedin-related proteins are extracellular matrix, secreted glycoproteins with conserved C-terminal motifs. They are expressed in a wide variety of tissues and in broad range of species, from *Caenorhabditis elegans* to *Homo sapiens*. Olfactomedin-related proteins comprise a gene family with at least 5 family members in humans. One of the five, TIGR/myocilin protein, is expressed in the eye and is associated with the pathogenesis of glaucoma (Kulkarni, N.H. et al. (2000) Genet. Res. 76:41-50). Research by Yokoyama et al. (1996) found a 135-amino acid protein, termed AMY, having 96% sequence identity with rat neuronal olfactomedin-releated ER localized protein in a neuroblastoma cell line cDNA library, suggesting an essential role for AMY in nerve tissue (Yokoyama, M. et al. (1996) DNA Res. 3:311-320). Neuron-specific olfactomedin-related glycoproteins isolated from rat brain cDNA libraries show strong sequence similarity with olfactomedin. This similarity is suggestive of a matrix-related function of these glycoproteins in neurons and neurosecretory cells (Danielson, P.E. et al. (1994) J. Neurosci. Res. 38:468-478).

Mac-2 binding protein is a 90-kD serum protein (90K), a secreted glycoprotein isolated from both the human breast carcinoma cell line SK-BR-3, and human breast milk. It specifically binds to a human macrophage-associated lectin, Mac-2. Structurally, the mature protein is 567 amino acids in length and is proceeded by an 18-amino acid leader. There are 16 cysteines and seven potential N-linked glycosylation sites. The first 106 amino acids represent a domain very similar to an ancient protein superfamily defined by a macrophage scavenger receptor cysteine-rich domain (Koths,K. et al. (1993) J. Biol. Chem. 268:14245-14249). 90K is elevated in the serum of subpopulations of AIDS patients and is expressed at varying levels in primary tumor samples and tumor cell lines. Ullrich et al. (1994) have demonstrated that 90K stimulates host defense systems and can induce interleukin-2 secretion. This immune stimulation is proposed to be a result of oncogenic transformation, viral infection or pathogenic invasion (Ullrich,A., et al. (1994) J. Biol. Chem. 269:18401-18407).

Semaphorins are a large group of axonal guidance molecules consisting of at least 30 different members and are found in vertebrates, invertebrates, and even certain viruses. All semaphorins contain the sema domain which is approximately 500 amino acids in length. Neuropilin, a semaphorin receptor, has been shown to promote neurite outgrowth in vitro. The extracellular region of neuropilins consists of three different domains: CUB, discoidin, and MAM domains. The CUB and the MAM motifs of neuropilin have been suggested to have roles in protein-protein interactions and are thought to be involved in the binding of semaphorins through the sema and the C-terminal domains (reviewed in Raper, J.A. (2000) Curr. Opin. Neurobiol. 10:88-94). Plexins are neuronal cell surface molecules that mediate cell adhesion via a homophilic binding mechanism in the presence of calcium ions. Plexins have been shown to be expressed in the receptors and neurons of

particular sensory systems (Ohta, K. et al. (1995) Cell 14:1189-1199). There is evidence that suggests that some plexins function to control motor and CNS axon guidance in the developing nervous system. Plexins, which themselves contain complete semaphorin domains, may be both the ancestors of classical semaphorins and binding partners for semaphorins (Winberg, M.L. et al (1998) Cell 95:903-916).

Human pregnancy-specific beta 1-glycoprotein (PSG) is a family of closely related glycoproteins of molecular weights of 72 KDa, 64KDa, 62KDa, and 54KDa. Together with the carcinoembryonic antigen, they comprise a subfamily within the immunoglobulin superfamily (Plouzek, C.A. and Chou, J.Y. (1991) Endocrinology 129:950-958) Different subpopulations of PSG have been found to be produced by the trophoblasts of the human placenta, and the amnionic and chorionic membranes (Plouzek, C.A. et al. (1993) Placenta 14:277-285).

Autocrine motility factor (AMF) is one of the motility cytokines regulating tumor cell migration; therefore identification of the signaling pathway coupled with it has critical importance. Autocrine motility factor receptor (AMFR) expression has been found to be associated with tumor progression in thymoma (Ohta Y. et al. (2000) Int. J. Oncol. 17:259-264). AMFR is a cell surface glycoprotein of molecular weight 78KDa.

Hormones are secreted molecules that travel through the circulation and bind to specific receptors on the surface of, or within, target cells. Although they have diverse biochemical compositions and mechanisms of action, hormones can be grouped into two categories. One category includes small lipophilic hormones that diffuse through the plasma membrane of target cells, bind to cytosolic or nuclear receptors, and form a complex that alters gene expression. Examples of these molecules include retinoic acid, thyroxine, and the cholesterol-derived steroid hormones such as progesterone, estrogen, testosterone, cortisol, and aldosterone. The second category includes hydrophilic hormones that function by binding to cell surface receptors that transduce signals across the plasma membrane. Examples of such hormones include amino acid derivatives such as catecholamines (epinephrine, norepinephrine) and histamine, and peptide hormones such as glucagon, insulin, gastrin, secretin, cholecystokinin, adrenocorticotropic hormone, follicle stimulating hormone, luteinizing hormone, thyroid stimulating hormone, and vasopressin. (See, for example, Lodish et al. (1995) Molecular Cell Biology, Scientific American Books Inc., New York, NY, pp. 856-864.)

Pro-opiomelanocortin (POMC) is the precursor polypeptide of corticotropin (ACTH), a hormone synthesized by the anterior pituitary gland, which functions in the stimulation of the adrenal cortex. POMC is also the precursor polypeptide of the hormone beta-lipotropin (beta-LPH). Each hormone includes smaller peptides with distinct biological activities: alpha-melanotropin (alpha-MSH) and corticotropin-like intermediate lobe peptide (CLIP) are formed from ACTH; gamma-lipotropin (gamma-LPH) and beta-endorphin are peptide components of beta-LPH; while beta-MSH

is contained within gamma-LPH. Adrenal insufficiency due to ACTH deficiency, resulting from a genetic mutation in exons 2 and 3 of POMC results in an endocrine disorder characterized by early-onset obesity, adrenal insufficiency, and red hair pigmentation (Chretien, M. et al. (1979) Canad. J. Biochem. 57:1111-1121; Krude, H. et al. (1998) Nature Genet. 19:155-157; Online Mendelian Inheritance in Man (OMIM) 176830).

Growth and differentiation factors are secreted proteins which function in intercellular communication. Some factors require oligomerization or association with membrane proteins for activity. Complex interactions among these factors and their receptors trigger intracellular signal transduction pathways that stimulate or inhibit cell division, cell differentiation, cell signaling, and cell motility. Most growth and differentiation factors act on cells in their local environment (paracrine signaling). There are three broad classes of growth and differentiation factors. The first class includes the large polypeptide growth factors such as epidermal growth factor, fibroblast growth factor, transforming growth factor, insulin-like growth factor, and platelet-derived growth factor. The second class includes the hematopoietic growth factors such as the colony stimulating factors (CSFs). Hematopoietic growth factors stimulate the proliferation and differentiation of blood cells such as B-lymphocytes, T-lymphocytes, erythrocytes, platelets, eosinophils, basophils, neutrophils, macrophages, and their stem cell precursors. The third class includes small peptide factors such as bombesin, vasopressin, oxytocin, endothelin, transferrin, angiotensin II, vasoactive intestinal peptide, and bradykinin, which function as hormones to regulate cellular functions other than proliferation.

Growth and differentiation factors play critical roles in neoplastic transformation of cells in vitro and in tumor progression in vivo. Inappropriate expression of growth factors by tumor cells may contribute to vascularization and metastasis of tumors. During hematopoiesis, growth factor misregulation can result in anemias, leukemias, and lymphomas. Certain growth factors such as interferon are cytotoxic to tumor cells both in vivo and in vitro. Moreover, some growth factors and growth factor receptors are related both structurally and functionally to oncoproteins. In addition, growth factors affect transcriptional regulation of both proto-oncogenes and oncosuppressor genes. (Reviewed in Pimentel, E. (1994) Handbook of Growth Factors, CRC Press, Ann Arbor, MI, pp. 1-9.)

The Slit protein, first identified in Drosophila, is critical in central nervous system midline formation and potentially in nervous tissue histogenesis and axonal pathfinding. Itoh et al. have identified mammalian homologues of the slit gene (human Slit-1, Slit-2, Slit-3 and rat Slit-1). The encoded proteins are putative secreted proteins containing EGF-like motifs and leucine-rich repeats, both of which are conserved protein-protein interaction domains. Slit-1, -2, and -3 mRNAs are expressed in the brain, spinal cord, and thyroid, respectively (Itoh, A. et al., (1998) Brain Res. Mol. Brain Res. 62:175-186). The Slit family of proteins are indicated to be functional ligands of glypican-1 in nervous tissue and suggests that their interactions may be critical in certain stages

during central nervous system histogenesis (Liang, Y. et al., (1999) J. Biol. Chem. 274:17885-17892).

Neuropeptides and vasomediators (NP/VM) comprise a large family of endogenous signaling molecules. Included in this family are neuropeptides and neuropeptide hormones such as bombesin, neuropeptide Y, neurotensin, neuromedin N, melanocortins, opioids, galanin, somatostatin, tachykinins, urotensin II and related peptides involved in smooth muscle stimulation, vasopressin, vasoactive intestinal peptide, and circulatory system-borne signaling molecules such as angiotensin, complement, calcitonin, endothelins, formyl-methionyl peptides, glucagon, cholecystokinin and gastrin. NP/VMs can transduce signals directly, modulate the activity or release of other neurotransmitters and hormones, and act as catalytic enzymes in cascades. The effects of NP/VMs range from extremely brief to long-lasting. (Reviewed in Martin, C.R. et al. (1985) Endocrine Physiology, Oxford University Press, New York, NY, pp. 57-62.)

NP/VMs are involved in numerous neurological and cardiovascular disorders. For example, neuropeptide Y is involved in hypertension, congestive heart failure, affective disorders, and appetite regulation. Somatostatin inhibits secretion of growth hormone and prolactin in the anterior pituitary, as well as inhibiting secretion in intestine, pancreatic acinar cells, and pancreatic beta-cells. A reduction in somatostatin levels has been reported in Alzheimer's disease and Parkinson's disease. Vasopressin acts in the kidney to increase water and sodium absorption, and in higher concentrations stimulates contraction of vascular smooth muscle, platelet activation, and glycogen breakdown in the liver. Vasopressin and its analogues are used clinically to treat diabetes insipidus. Endothelin and angiotensin are involved in hypertension, and drugs, such as captopril, which reduce plasma levels of angiotensin, are used to reduce blood pressure (Watson, S. and S. Arkinstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego CA, pp. 194; 252; 284; 55; 111).

Neuropeptides have also been shown to have roles in nociception (pain). Vasoactive intestinal peptide appears to play an important role in chronic neuropathic pain. Nociceptin, an endogenous ligand for for the opioid receptor-like 1 receptor, is thought to have a predominantly anti-nociceptive effect, and has been shown to have analgesic properties in different animal models of tonic or chronic pain (Dickinson, T. and Fleetwood-Walker, S.M. (1998) Trends Pharmacol. Sci. 19:346-348).

Other proteins that contain signal peptides include secreted proteins with enzymatic activity, or enzyme inhibitory activity. Such activity includes, for example, oxidoreductase/dehydrogenase activity, transferase activity, hydrolase activity, lyase activity, isomerase activity, or ligase activity. For example, matrix metalloproteinases are secreted hydrolytic enzymes that degrade the extracellular matrix and thus play an important role in tumor metastasis, tissue morphogenesis, and arthritis (Reponen, P. et al. (1995) Dev. Dyn. 202:388-396; Firestein, G.S. (1992) Curr. Opin. Rheumatol. 4:348-354; Ray, J.M. and Stetler-Stevenson, W.G. (1994) Eur. Respir. J. 7:2062-2072; and Mignatti,

P. and Rifkin, D.B. (1993) Physiol. Rev. 73:161-195). Tissue Inhibitors of Metalloproteinase (TIMPs), on the other hand, are secreted proteins which bind to metalloproteinases and block their activity (Stetler-Stevenson, W.G. et al. (1989) J. Biol. Chem. 264:17374-17378). Additional examples are the acetyl-CoA synthetases which activate acetate for use in lipid synthesis or energy generation (Luong, A. et al. (2000) J. Biol. Chem. 275:26458-26466). The result of acetyl-CoA synthetase activity is the formation of acetyl-CoA from acetate and CoA. Acetyl-CoA sythetases share a region of sequence similarity identified as the AMP-binding domain signature. Acetyl-CoA synthetase has been shown to be associated with hypertension (H. Toh (1991) Protein Seq. Data Anal. 4:111-117; and Iwai, N. et al., (1994) Hypertension 23:375-380).

A number of isomerases catalyze steps in protein folding, phototransduction, and various anabolic and catabolic pathways. One class of isomerases is known as peptidyl-prolyl *cis-trans* isomerases (PPIases). PPIases catalyze the *cis* to *trans* isomerization of certain proline imidic bonds in proteins. Two families of PPIases are the FK506 binding proteins (FKBPs), and cyclophilins (CyPs). FKBPs bind the potent immunosuppressants FK506 and rapamycin, thereby inhibiting signaling pathways in T-cells. Specifically, the PPIase activity of FKBPs is inhibited by binding of FK506 or rapamycin. There are five members of the FKBP family which are named according to their calculated molecular masses (FKBP12, FKBP13, FKBP25, FKBP52, and FKBP65), and localized to different regions of the cell where they associate with different protein complexes (Coss, M. et al. (1995) J. Biol. Chem. 270:29336 - 29341; Schreiber, S.L. (1991) Science 251:283 - 287).

The peptidyl-prolyl isomerase activity of CyP may be part of the signaling pathway that leads to T-cell activation. CyP isomerase activity is associated with protein folding and protein trafficking, and may also be involved in assembly/disassembly of protein complexes and regulation of protein activity. For example, in *Drosophila*, the CyP NinaA is required for correct localization of rhodopsins, while a mammalian CyP (Cyp40) is part of the Hsp90/Hsc70 complex that binds steroid receptors. The mammalian CypA has been shown to bind the *gag* protein from human immunodeficiency virus 1 (HIV-1), an interaction that can be inhibited by cyclosporin. Since cyclosporin has potent anti-HIV-1 activity, CypA may play an essential function in HIV-1 replication. Finally, Cyp40 has been shown to bind and inactivate the transcription factor c-Myb, an effect that is reversed by cyclosporin. This effect implicates CyPs in the regulation of transcription, transformation, and differentiation (Bergsma, D.J. et al (1991) J. Biol. Chem. 266:23204 - 23214; Hunter, T. (1998) Cell 92: 141-143; and Leverson, J.D. and Ness, S.A. (1998) Mol. Cell. 1:203-211).

Another protein that contains a signal peptide is encoded by the seizure-related gene, SEZ-6, a brain specific cDNA whose expression is increased by the convulsant drug pentylentetrazole. The SEZ-6 protein is expressed in the cerebrum and cerebellum. SEZ-6 contains five short consensus repeats (SCR, or sushi domains) and two CUB (complement C1r/s-like repeat) domains in addition to

a signal peptide and a single transmembrane domain (Shimizu-Nishikawa, K. et al. (1995) Biochem. Biophys. Res. Commun. 216:382-389).

Gamma-carboxyglutamic acid (Gla) proteins rich in proline (PRGPs) are members of a family of vitamin K-dependent single-pass integral membrane proteins. These proteins are characterized by an extracellular amino terminal domain of approximately 45 amino acids rich in Gla. The intracellular carboxyl terminal region contains one or two copies of the sequence PPXY, a motif present in a variety of proteins involved in such diverse cellular functions as signal transduction, cell cycle progression, and protein turnover (Kulman, J.D. et al., (2001) Proc. Natl. Acad. Sci. U.S.A. 98:1370-1375). The process of post-translational modification of glutamic residues to form Gla is Vitamin K-dependent carboxylation. Proteins which contain Gla include plasma proteins involved in blood coagulation. These proteins are prothrombin, proteins C, S, and Z, and coagulation factors VII, IX, and X. Osteocalcin (bone-Gla protein, BGP) and matrix Gla-protein (MGP) also contain Gla (Friedman, P.A., and C.T. Przysiecki (1987) Int. J. Biochem. 19:1-7; C. Vermeer (1990) Biochem. J. 266:625-636).

The Drosophila sp. gene *crossveinless 2* is characterized as having a putative signal or transmembrane sequence, and a partial Von Willebrand Factor D domain similar to those domains known to regulate the formation of intramolecular and intermolecular bonds and five cysteine-rich domains, known to bind BMP-like (bone morphogenetic proteins) ligands. These features suggest that *crossveinless 2* may act extracelluarly or in the secretory pathway to directly potentiate ligand signaling and hence, involvement in the BMP-like signaling pathway known to play a role in vein specification (Conley, C.A. et al., (2000) Development 127:3947-3959). The dorsal-ventral patterning in both vertebrate and Drosophila embryos requires a conserved system of extracellular proteins to generate a positional informational gradient.

Immunoglobulins

Antigen recognition molecules are key players in the sophisticated and complex immune systems which all vertebrates have developed to provide protection from viral, bacterial, fungal, and parasitic infections. A key feature of the immune system is its ability to distinguish foreign molecules, or antigens, from "self" molecules. Most cell surface and soluble molecules that mediate functions such as recognition, adhesion or binding have evolved from a common evolutionary precursor (i.e., these proteins have structural homology). A number of molecules outside the immune system that have similar functions are also derived from this same evolutionary precursor. This ability is mediated primarily by secreted and transmembrane proteins expressed by leukocytes (white blood cells) such as lymphocytes, granulocytes, and monocytes. Most of these proteins belong to the immunoglobulin (Ig) superfamily, members of which contain one or more repeats of a conserved structural domain. This Ig domain is comprised of antiparallel β sheets joined by a disulfide bond in

an arrangement called the Ig fold. The criteria for a protein to be a member of the Ig superfamily is to have one or more Ig domains, which are regions of 70-110 amino acid residues in length homologous to either Ig variable-like (V) or Ig constant-like (C) domains. Members of the Ig superfamily include antibodies (Ab), T cell receptors (TCRs), class I and II major histocompatibility (MHC) proteins and immune cell-specific surface markers such as the "cluster of differentiation" or CD antigens, CD2, CD3, CD4, CD8, poly-Ig receptors, Fc receptors, neural cell-adhesion molecule (NCAM) and platelet-derived growth factor receptor (PDGFR). These antigens have been identified using systematic, monoclonal antibody (mAb)-based "shot gun" techniques. These techniques have resulted in the production of hundreds of mAbs directed against unknown cell surface leukocytic antigens. These antigens have been grouped into "clusters of differentiation" based on common immunocytochemical localization patterns in various differentiated and undifferentiated leukocytic cell types. Antigens in a given cluster are presumed to identify a single cell surface protein and are assigned a "cluster of differentiation" or "CD" designation. Some of the genes encoding proteins identified by CD antigens have been cloned and verified by standard molecular biology techniques. CD antigens have been characterized as both transmembrane proteins and cell surface proteins anchored to the plasma membrane via covalent attachment to fatty acid-containing glycolipids such as glycosylphosphatidylinositol (GPI). (Reviewed in Barclay, A. N. et al. (1995) The Leucocyte Antigen Facts Book, Academic Press, San Diego, CA, pp. 17-20.)

Ig domains (V and C) are regions of conserved amino acid residues that give a polypeptide a globular tertiary structure called an immunoglobulin (or antibody) fold, which consists of two approximately parallel layers of β-sheets. Conserved cysteine residues form an intrachain disulfide-bonded loop, 55-75 amino acid residues in length, which connects the two layers of the β-sheets. Each β-sheet has three or four anti-parallel β-strands of 5-10 amino acid residues. Hydrophobic and hydrophilic interactions of amino acid residues within the β-strands stabilize the Ig fold (hydrophobic on inward facing amino acid residues and hydrophilic on the amino acid residues in the outward facing portion of the strands). A V domain consists of a longer polypeptide than a C domain, with an additional pair of β-strands in the Ig fold.

A consistent feature of Ig superfamily genes is that each sequence of an Ig domain is encoded by a single exon. It is possible that the superfamily evolved from a gene coding for a single Ig domain involved in mediating cell-cell interactions. New members of the superfamily then arose by exon and gene duplications. Modern Ig superfamily proteins contain different numbers of V and/or C domains. Another evolutionary feature of this superfamily is the ability to undergo DNA rearrangements, a unique feature retained by the antigen receptor members of the family.

Many members of the Ig superfamily are integral plasma membrane proteins with extracellular Ig domains. The hydrophobic amino acid residues of their transmembrane domains and

their cytoplasmic tails are very diverse, with little or no homology among Ig family members or to known signal-transducing structures. There are exceptions to this general superfamily description. For example, the cytoplasmic tail of PDGFR has tyrosine kinase activity. In addition Thy-1 is a glycoprotein found on thymocytes and T cells. This protein has no cytoplasmic tail, but is instead attached to the plasma membrane by a covalent glycophosphatidylinositol linkage.

Another common feature of many Ig superfamily proteins is the interactions between Ig domains which are essential for the function of these molecules. Interactions between Ig domains of a multimeric protein can be either homophilic or heterophilic (i.e., between the same or different Ig domains). Antibodies are multimeric proteins which have both homophilic and heterophilic interactions between Ig domains. Pairing of constant regions of heavy chains forms the Fc region of an antibody and pairing of variable regions of light and heavy chains form the antigen binding site of an antibody. Heterophilic interactions also occur between Ig domains of different molecules. These interactions provide adhesion between cells for significant cell-cell interactions in the immune system and in the developing and mature nervous system. (Reviewed in Abbas, A.K. et al. (1991) Cellular and Molecular Immunology, W.B. Saunders Company, Philadelphia, PA, pp.142-145.)

Antibodies

MHC proteins are cell surface markers that bind to and present foreign antigens to T cells. MHC molecules are classified as either class I or class II. Class I MHC molecules (MHC I) are expressed on the surface of almost all cells and are involved in the presentation of antigen to cytotoxic T cells. For example, a cell infected with virus will degrade intracellular viral proteins and express the protein fragments bound to MHC I molecules on the cell surface. The MHC I/antigen complex is recognized by cytotoxic T-cells which destroy the infected cell and the virus within. Class II MHC molecules are expressed primarily on specialized antigen-presenting cells of the immune system, such as B-cells and macrophages. These cells ingest foreign proteins from the extracellular fluid and express MHC II/antigen complex on the cell surface. This complex activates helper T-cells, which then secrete cytokines and other factors that stimulate the immune response. MHC molecules also play an important role in organ rejection following transplantation. Rejection occurs when the recipient's T-cells respond to foreign MHC molecules on the transplanted organ in the same way as to self MHC molecules bound to foreign antigen. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of the Cell, Garland Publishing, New York, NY, pp. 1229-1246

Antibodies are multimeric members of the Ig superfamily which are either expressed on the surface of B-cells or secreted by B-cells into the circulation. Antibodies bind and neutralize foreign antigens in the blood and other extracellular fluids. The prototypical antibody is a tetramer consisting of two identical heavy polypeptide chains (H-chains) and two identical light polypeptide chains (L-chains) interlinked by disulfide bonds. This arrangement confers the characteristic Y-shape to

antibody molecules. Antibodies are classified based on their H-chain composition. The five antibody classes, IgA, IgD, IgE, IgG and IgM, are defined by the α, δ, ϵ, γ, and μ H-chain types. There are two types of L-chains, κ and λ, either of which may associate as a pair with any H-chain pair. IgG, the most common class of antibody found in the circulation, is tetrameric, while the other classes of antibodies are generally variants or multimers of this basic structure.

H-chains and L-chains each contain an N-terminal variable region and a C-terminal constant region. The constant region consists of about 110 amino acids in L-chains and about 330 or 440 amino acids in H-chains. The amino acid sequence of the constant region is nearly identical among H- or L-chains of a particular class. The variable region consists of about 110 amino acids in both H- and L-chains. However, the amino acid sequence of the variable region differs among H- or L-chains of a particular class. Within each H- or L-chain variable region are three hypervariable regions of extensive sequence diversity, each consisting of about 5 to 10 amino acids. In the antibody molecule, the H- and L-chain hypervariable regions come together to form the antigen recognition site. (Reviewed in Alberts, B. et al. supra, pp. 1206-1213 and 1216-1217.)

Both H-chains and L-chains contain the repeated Ig domains of members of the Ig superfamily. For example, a typical H-chain contains four Ig domains, three of which occur within the constant region and one of which occurs within the variable region and contributes to the formation of the antigen recognition site. Likewise, a typical L-chain contains two Ig domains, one of which occurs within the constant region and one of which occurs within the variable region.

The immune system is capable of recognizing and responding to any foreign molecule that enters the body. Therefore, the immune system must be armed with a full repertoire of antibodies against all potential antigens. Such antibody diversity is generated by somatic rearrangement of gene segments encoding variable and constant regions. These gene segments are joined together by site-specific recombination which occurs between highly conserved DNA sequences that flank each gene segment. Because there are hundreds of different gene segments, millions of unique genes can be generated combinatorially. In addition, imprecise joining of these segments and an unusually high rate of somatic mutation within these segments further contribute to the generation of a diverse antibody population.

The discovery of new secreted proteins, and the polynucleotides encoding them, satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of secreted proteins.

SUMMARY OF THE INVENTION

The invention features purified polypeptides, secreted proteins, referred to collectively as "SECP" and individually as "SECP-1," "SECP-2," "SECP-3," "SECP-4," "SECP-5," "SECP-6," "SECP-7," "SECP-8," "SECP-9," "SECP-10," "SECP-11," "SECP-12," "SECP-13," "SECP-14," "SECP-15," "SECP-16," "SECP-17," "SECP-18," "SECP-19," "SECP-20," "SECP-21," "SECP-22," "SECP-23," "SECP-24," "SECP-25," "SECP-26," "SECP-27," "SECP-28," "SECP-29," "SECP-30," "SECP-31," "SECP-32," "SECP-33," "SECP-34," "SECP-35," "SECP-36," "SECP-37," "SECP-38," "SECP-39," "SECP-40," "SECP-41," "SECP-42," "SECP-43," "SECP-44," "SECP-45," "SECP-46," "SECP-47," "SECP-48," "SECP-49," "SECP-50," "SECP-51," "SECP-52," "SECP-53," "SECP-54," "SECP-55," "SECP-56," "SECP-57," "SECP-58," "SECP-59," "SECP-60," "SECP-61," "SECP-62," "SECP-63," "SECP-64," "SECP-65," "SECP-66," and "SECP-67." In one aspect, the invention provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-67.

The invention further provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-67. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:68-134.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, and d) an immunogenic fragment of a polypeptide having an amino

acid sequence selected from the group consisting of SEQ ID NO:1-67. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67.

The invention further provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:68-134, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:68-134, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:68-134, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:68-134, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The

method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:68-134, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:68-134, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-67. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, and d) an

immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional SECP, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, c) a biologically active fragment of a polypeptide

having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:68-134, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:68-134, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:68-134, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:68-134, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:68-134, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the present invention.

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog for polypeptides of the invention. The probability score for the match between each polypeptide and its GenBank homolog is also shown.

Table 3 shows structural features of polypeptide sequences of the invention, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of the polypeptides.

Table 4 lists the cDNA and/or genomic DNA fragments which were used to assemble polynucleotide sequences of the invention, along with selected fragments of the polynucleotide sequences.

Table 5 shows the representative cDNA library for polynucleotides of the invention.

Table 6 provides an appendix which describes the tissues and vectors used for construction of the cDNA libraries shown in Table 5.

Table 7 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing

the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"SECP" refers to the amino acid sequences of substantially purified SECP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of SECP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of SECP either by directly interacting with SECP or by acting on components of the biological pathway in which SECP participates.

An "allelic variant" is an alternative form of the gene encoding SECP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding SECP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as SECP or a polypeptide with at least one functional characteristic of SECP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding SECP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding SECP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent SECP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of SECP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide,

polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of SECP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of SECP either by directly interacting with SECP or by acting on components of the biological pathway in which SECP participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind SECP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "aptamer" refers to a nucleic acid or oligonucleotide molecule that binds to a specific molecular target. Aptamers are derived from an in vitro evolutionary process (e.g., SELEX (Systematic Evolution of Ligands by EXponential Enrichment), described in U.S. Patent No. 5,270,163), which selects for target-specific aptamer sequences from large combinatorial libraries. Aptamer compositions may be double-stranded or single-stranded, and may include deoxyribonucleotides, ribonucleotides, nucleotide derivatives, or other nucleotide-like molecules. The nucleotide components of an aptamer may have modified sugar groups (e.g., the 2'-OH group of a ribonucleotide may be replaced by 2'-F or 2'-NH_2), which may improve a desired property, e.g., resistance to nucleases or longer lifetime in blood. Aptamers may be conjugated to other molecules,

e.g., a high molecular weight carrier to slow clearance of the aptamer from the circulatory system. Aptamers may be specifically cross-linked to their cognate ligands, e.g., by photo-activation of a cross-linker. (See, e.g., Brody, E.N. and L. Gold (2000) J. Biotechnol. 74:5-13.)

The term "intramer" refers to an aptamer which is expressed in vivo. For example, a vaccinia virus-based RNA expression system has been used to express specific RNA aptamers at high levels in the cytoplasm of leukocytes (Blind, M. et al. (1999) Proc. Natl Acad. Sci. USA 96:3606-3610).

The term "spiegelmer" refers to an aptamer which includes L-DNA, L-RNA, or other left-handed nucleotide derivatives or nucleotide-like molecules. Aptamers containing left-handed nucleotides are resistant to degradation by naturally occurring enzymes, which normally act on substrates containing right-handed nucleotides.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic SECP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'. In an alternative example, SEQ ID NO:135 and SEQ ID NO:136 pair with their complements, SEQ ID NO:114 and SEQ ID NO:116, respectively.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding SECP or fragments of SECP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be

associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

"Differential expression" refers to increased or upregulated; or decreased, downregulated, or absent gene or protein expression, determined by comparing at least two different samples. Such comparisons may be carried out between, for example, a treated and an untreated sample, or a diseased and a normal sample.

"Exon shuffling" refers to the recombination of different coding regions (exons). Since an exon may represent a structural or functional domain of the encoded protein, new proteins may be assembled through the novel reassortment of stable substructures, thus allowing acceleration of the evolution of new protein functions.

A "fragment" is a unique portion of SECP or the polynucleotide encoding SECP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:68-134 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:68-134, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:68-134 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:68-134 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:68-134 and the region of SEQ ID NO:68-134 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-67 is encoded by a fragment of SEQ ID NO:68-134. A fragment of SEQ ID NO:1-67 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-67. For example, a fragment of SEQ ID NO:1-67 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-67. The precise length of a fragment of SEQ ID NO:1-67 and the region of SEQ ID NO:1-67 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full length" polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full length" polynucleotide sequence encodes a "full length" polypeptide sequence.

"Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html.

The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by

CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about

1% (w/v) SDS, and about 100 μg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 μg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of SECP which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment

of SECP which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of SECP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of SECP.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an SECP may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of SECP.

"Probe" refers to nucleic acid sequences encoding SECP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al. (1987) Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al. (1990) PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of

oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing SECP, nucleic acids encoding SECP, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A

and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed cells" includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human secreted proteins (SECP), the polynucleotides encoding SECP, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders.

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an Incyte

polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide sequence is denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown.

Table 2 shows sequences with homology to the polypeptides of SEQ ID NO:1-7, SEQ ID NO:34-35, and SEQ ID NO:57-58, as identified by BLAST analysis against the GenBank protein (genpept) database. Columns 1 and 2 show polypeptides of SEQ ID NO:1-7, SEQ ID NO:34-35, and SEQ ID NO:57-58 and their corresponding Incyte polypeptide sequence numbers (Incyte Polypeptide ID). Column 3 shows the GenBank identification number (Genbank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability score for the match between SEQ ID NO:1-7, SEQ ID NO:34-35, and SEQ ID NO:57-58 and their GenBank homologs. Column 5 shows the annotation of the GenBank homolog along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

Table 3 shows various structural features of each of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Genetics Computer Group, Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs, including the locations of signal peptides (as indicated by "Signal Peptide" and/or "signal_cleavage"). Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are secreted proteins. For example, SEQ ID NO:1 is 56% identical to a human cerebral cell adhesion molecule (GenBank ID g5764665) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 3.8e-156, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:1 also contains a lysyl hydrolase domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from SPSCAN, HMMER, BLAST_PRODOM and BLAST_DOMO analyses using other sequence databases provide further corroborative evidence that SEQ ID NO:1 is a secreted hydrolase. In an alternative example, SEQ ID NO:2 is 32% identical to mouse seizure-related gene product 6 precursor (GenBank ID g693910) and is 67% identical from residue S22 to residue R527 to human CUB and sushi multiple domains 1

protein (GenBank ID g14794726) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability scores are 2.9e-42 and 0.0 respectively, which indicate the probabilities of obtaining the observed polypeptide sequence alignments by chance. SEQ ID NO:2 also contains three sushi domains and two CUB domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) In addition, SEQ ID NO:2 contains a signal peptide as identified by HMMER analysis. Data from BLIMPS analysis provides further corroborative evidence that SEQ ID NO:2 is a secreted protein which contains sushi domains. In an alternative example, SEQ ID NO:3 shares 51% local identity to a mouse transmembrane protein (GenBank ID g7259265) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 2.4e-28, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:3 also contains a signal peptide as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from SPSCAN analyses provide further corroborative evidence that SEQ ID NO:3 is a secreted protein. In an alternative example, SEQ ID NO:58 is 39% identical to ZOG, a rat zona glomerulosa specific protein (GenBank ID g3097285) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 1.2e-65, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:58 contains a signal peptide and single transmembrane domain. SEQ ID NO:58 also contains a number of EGF-like domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) The presence of this motif is confirmed by BLIMPS and MOTIFS analyses, providing further corroborative evidence that SEQ ID NO:58 is a secreted protein. SEQ ID NO:4-57 and SEQ ID NO:59-67 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-67 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention. Column 3 shows the length of each polynucleotide sequence in base pairs. Column 4 lists fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:68-136 or that distinguish between SEQ ID NO:68-136 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA

sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages comprised of both cDNA and genomic DNA. These sequences were used to assemble the full length polynucleotide sequences of the invention. Columns 6 and 7 of Table 4 show the nucleotide start (5′) and stop (3′) positions of the cDNA and/or genomic sequences in column 5 relative to their respective full length sequences.

The identification numbers in Column 5 of Table 4 may refer specifically, for example, to Incyte cDNAs along with their corresponding cDNA libraries. For example, 8052177J1 is the identification number of an Incyte cDNA sequence, and FTUBTUE01 is the cDNA library from which it is derived. Incyte cDNAs for which cDNA libraries are not indicated were derived from pooled cDNA libraries (e.g., 71926854V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g2204647) which contributed to the assembly of the full length polynucleotide sequences. In addition, the identification numbers in column 5 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the identification numbers in column 5 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation "NP"). Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, FL_*XXXXXX*_N_1_N_2_*YYYYY*_N_3_N_4 represents a "stitched" sequence in which *XXXXXX* is the identification number of the cluster of sequences to which the algorithm was applied, and *YYYYY* is the number of the prediction generated by the algorithm, and $N_{1,2,3...}$, if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the identification numbers in column 5 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, FL*XXXXXX*_g*AAAAA*_g*BBBBB*_1_*N* is the identification number of a "stretched" sequence, with *XXXXXX* being the Incyte project identification number, g*AAAAA* being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, g*BBBBB* being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and *N* referring to specific exons (See Example V). In instances where a RefSeq sequence was used as a protein homolog for the "exon-stretching" algorithm, a RefSeq identifier (denoted by "NM," "NP," or "NT") may be used in place of the GenBank identifier (*i.e.*, g*BBBBB*).

Alternatively, a prefix identifies component sequences that were hand-edited, predicted from genomic DNA sequences, or derived from a combination of sequence analysis methods. The following Table lists examples of component sequence prefixes and corresponding sequence analysis

methods associated with the prefixes (see Example IV and Example V).

Prefix	Type of analysis and/or examples of programs
GNN, GFG, ENST	Exon prediction from genomic sequences using, for example, GENSCAN (Stanford University, CA, USA) or FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK).
GBI	Hand-edited analysis of genomic sequences.
FL	Stitched or stretched genomic sequences (see Example V).
INCY	Full length transcript and exon prediction from mapping of EST sequences to the genome. Genomic location and EST composition data are combined to predict the exons and resulting transcript.

In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in column 5 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

Table 5 shows the representative cDNA libraries for those full length polynucleotide sequences which were assembled using Incyte cDNA sequences. The representative cDNA library is the Incyte cDNA library which is most frequently represented by the Incyte cDNA sequences which were used to assemble and confirm the above polynucleotide sequences. The tissues and vectors which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

The invention also encompasses SECP variants. A preferred SECP variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the SECP amino acid sequence, and which contains at least one functional or structural characteristic of SECP.

The invention also encompasses polynucleotides which encode SECP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:68-134, which encodes SECP. The polynucleotide sequences of SEQ ID NO:68-134, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding SECP. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding SECP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:68-134 which has at least about 70%, or alternatively at least about 85%, or even at least about

95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:68-134. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of SECP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding SECP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring SECP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode SECP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring SECP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding SECP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding SECP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode SECP and SECP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding SECP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:68-134 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied

Biosystems), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding SECP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T)

library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode SECP may be cloned in recombinant DNA molecules that direct expression of SECP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express SECP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter SECP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent No. 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of SECP, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are

optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding SECP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, SECP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of SECP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, supra, pp. 28-53.)

In order to express a biologically active SECP, the nucleotide sequences encoding SECP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding SECP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding SECP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding SECP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding SECP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding SECP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding SECP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding SECP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding SECP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of SECP are needed, e.g. for the production of antibodies, vectors which direct high level expression of SECP may be used. For example, vectors

containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of SECP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of SECP. Transcription of sequences encoding SECP may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding SECP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses SECP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of SECP in cell lines is preferred. For example, sequences encoding SECP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media

before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*⁻ and *apr*⁻ cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding SECP is inserted within a marker gene sequence, transformed cells containing sequences encoding SECP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding SECP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding SECP and that express SECP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of SECP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing

monoclonal antibodies reactive to two non-interfering epitopes on SECP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding SECP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding SECP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding SECP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode SECP may be designed to contain signal sequences which direct secretion of SECP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding SECP may be ligated to a heterologous sequence resulting in translation of a

fusion protein in any of the aforementioned host systems. For example, a chimeric SECP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of SECP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the SECP encoding sequence and the heterologous protein sequence, so that SECP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled SECP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ^{35}S-methionine.

SECP of the present invention or fragments thereof may be used to screen for compounds that specifically bind to SECP. At least one and up to a plurality of test compounds may be screened for specific binding to SECP. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of SECP, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which SECP binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express SECP, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing SECP or cell membrane fractions which contain SECP are then contacted with a test compound and binding, stimulation, or inhibition of activity of either SECP or the

compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with SECP, either in solution or affixed to a solid support, and detecting the binding of SECP to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

SECP of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of SECP. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for SECP activity, wherein SECP is combined with at least one test compound, and the activity of SECP in the presence of a test compound is compared with the activity of SECP in the absence of the test compound. A change in the activity of SECP in the presence of the test compound is indicative of a compound that modulates the activity of SECP. Alternatively, a test compound is combined with an in vitro or cell-free system comprising SECP under conditions suitable for SECP activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of SECP may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding SECP or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential

therapeutic or toxic agents.

Polynucleotides encoding SECP may also be manipulated in vitro in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding SECP can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding SECP is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress SECP, e.g., by secreting SECP in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of SECP and secreted proteins. In addition, the expression of SECP is closely associated with neurological, gastrointestinal, cardiovascular, reproductive, developmental, diseased, and tumorous tissues such as adrenal gland tumor tissue. Therefore, SECP appears to play a role in cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders. In the treatment of disorders associated with increased SECP expression or activity, it is desirable to decrease the expression or activity of SECP. In the treatment of disorders associated with decreased SECP expression or activity, it is desirable to increase the expression or activity of SECP.

Therefore, in one embodiment, SECP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SECP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory

distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a cardiovascular disorder such as congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, complications of cardiac transplantation, arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental

disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; and a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Syndenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss.

In another embodiment, a vector capable of expressing SECP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SECP including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified SECP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SECP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of SECP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SECP including, but not limited to, those listed above.

In a further embodiment, an antagonist of SECP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of SECP. Examples of such disorders include, but are not limited to, those cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders described above. In one aspect, an antibody which specifically binds SECP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express SECP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding SECP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of SECP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The

combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of SECP may be produced using methods which are generally known in the art. In particular, purified SECP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind SECP. Antibodies to SECP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with SECP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to SECP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of SECP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to SECP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce SECP-specific single

chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for SECP may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between SECP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering SECP epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for SECP. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of SECP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple SECP epitopes, represents the average affinity, or avidity, of the antibodies for SECP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular SECP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the SECP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of SECP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of SECP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding SECP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding SECP. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding SECP. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) J. Allergy Clin. Immunol. 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) Blood 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) Pharmacol. Ther. 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) Br. Med. Bull. 51(1):217-225; Boado, R.J. et al. (1998) J. Pharm. Sci. 87(11):1308-1315; and Morris, M.C. et al. (1997) Nucleic Acids Res. 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding SECP may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassamias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal,

R.G. (1995) Science 270:404-410; Verma, I.M. and N. Somia (1997) Nature 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA. 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in SECP expression or regulation causes disease, the expression of SECP from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in SECP are treated by constructing mammalian expression vectors encoding SECP and introducing these vectors by mechanical means into SECP-deficient cells. Mechanical transfer technologies for use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J-L. and H. Récipon (1998) Curr. Opin. Biotechnol. 9:445-450).

Expression vectors that may be effective for the expression of SECP include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX, PCR2-TOPOTA vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). SECP may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β-actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. USA 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and Blau, H.M. supra)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding SECP from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method

(Graham, F.L. and A.J. Eb (1973) Virology 52:456-467), or by electroporation (Neumann, E. et al. (1982) EMBO J. 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to SECP expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding SECP under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent No. 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., $CD4^+$ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding SECP to cells which have one or more genetic abnormalities with respect to the expression of SECP. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent No. 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544 and Verma, I.M. and N. Somia (1997) Nature 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver

polynucleotides encoding SECP to target cells which have one or more genetic abnormalities with respect to the expression of SECP. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing SECP to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res. 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent No. 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent No. 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) J. Virol. 73:519-532 and Xu, H. et al. (1994) Dev. Biol. 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding SECP to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotechnol. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for SECP into the alphavirus genome in place of the capsid-coding region results in the production of a large number of SECP-coding RNAs and the synthesis of high levels of SECP in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of SECP into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA

transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding SECP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding SECP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs

and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding SECP. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased SECP expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding SECP may be therapeutically useful, and in the treatment of disorders associated with decreased SECP expression or activity, a compound which specifically promotes expression of the polynucleotide encoding SECP may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding SECP is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an in vitro cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding SECP are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding SECP. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a Schizosaccharomyces pombe gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res.

28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of SECP, antibodies to SECP, and mimetics, agonists, antagonists, or inhibitors of SECP.

The compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination

of an effective dose is well within the capability of those skilled in the art.

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising SECP or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, SECP or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example SECP or fragments thereof, antibodies of SECP, and agonists, antagonists or inhibitors of SECP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art.

Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind SECP may be used for the diagnosis of disorders characterized by expression of SECP, or in assays to monitor patients being treated with SECP or agonists, antagonists, or inhibitors of SECP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for SECP include methods which utilize the antibody and a label to detect SECP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring SECP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of SECP expression. Normal or standard values for SECP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to SECP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of SECP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding SECP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of SECP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of SECP, and to monitor regulation of SECP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding SECP or closely related molecules may be used to identify nucleic acid sequences which encode SECP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding SECP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the SECP encoding sequences. The hybridization probes of the subject

invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:68-134 or from genomic sequences including promoters, enhancers, and introns of the SECP gene.

Means for producing specific hybridization probes for DNAs encoding SECP include the cloning of polynucleotide sequences encoding SECP or SECP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding SECP may be used for the diagnosis of disorders associated with expression of SECP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a cardiovascular disorder such as congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic

endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, complications of cardiac transplantation, arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; and a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Syndenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss. The polynucleotide sequences encoding SECP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered SECP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding SECP may be useful in assays that

detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding SECP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding SECP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of SECP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding SECP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding SECP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding SECP, or a fragment of a polynucleotide complementary to the polynucleotide encoding SECP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or

quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding SECP may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding SECP are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of SECP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and

effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, SECP, fragments of SECP, or antibodies specific for SECP may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression in vivo, as in the case of a tissue or biopsy sample, or in vitro, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms,

knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at http://www.niehs.nih.gov/oc/news/toxchip.htm.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, supra). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for SECP to quantify the levels of SECP expression. In one embodiment, the antibodies are used as elements on a microarray,

and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) Anal. Biochem. 270:103-111; Mendoze, L.G. et al. (1999) Biotechniques 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) Electrophoresis 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed.

(1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding SECP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding SECP on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, SECP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a

solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between SECP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with SECP, or fragments thereof, and washed. Bound SECP is then detected by methods well known in the art. Purified SECP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding SECP specifically compete with a test compound for binding SECP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with SECP.

In additional embodiments, the nucleotide sequences which encode SECP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, including U.S. Ser. No. 60/236,869, U.S. Ser. No. 60/240,108, U.S. Ser. No. 60/239,812, U.S. Ser. No. 60/241,282, and U.S. Ser. No. 60/242,218, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA) and shown in Table 4, column 5. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), PCR2-TOPOTA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), pIGEN (Incyte Genomics, Palo Alto CA), or pINCY (Incyte Genomics), or derivatives thereof. Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence

scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and hidden Markov model (HMM)-based protein family databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide of the invention may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite,

and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:68-136. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 2.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative secreted proteins were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpri and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) J. Mol. Biol. 268:78-94, and Burge, C. and S. Karlin (1998) Curr. Opin. Struct. Biol. 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode secreted proteins, the encoded polypeptides were analyzed by querying against PFAM models for secreted proteins. Potential secreted proteins were also identified by homology to Incyte cDNA sequences that had been annotated as secreted proteins. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpri public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA

coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

"Stitched" Sequences

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpri public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either Incyte cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The

GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences were therefore "stretched" or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

VI. Chromosomal Mapping of SECP Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:68-134 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:68-134 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (http://www.ncbi.nlm.nih.gov/genemap/), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

In this manner, SEQ ID NO:70 was mapped to chromosome 5 within the interval from 79.2 to 92.3 centiMorgans. SEQ ID NO:98 was mapped to chromosome 4 within the interval from 145.3 to 146.4 centiMorgans.

VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel (1995) supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the

computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score x Percent Identity}}{\text{5 x minimum \{length(Seq. 1), length(Seq. 2)\}}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

Alternatively, polynucleotide sequences encoding SECP are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding SECP. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

VIII. Extension of SECP Encoding Polynucleotides

Full length polynucleotide sequences were also produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 µl to 10 µl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site

overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, full length polynucleotide sequences are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:68-136 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

X. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing

photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra.), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schena (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorbtion and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)$^+$ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)$^+$ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/μl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/μl RNase inhibitor, 500 μM dATP, 500 μM dGTP, 500 μM dTTP, 40 μM dCTP, 40 μM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)$^+$ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)$^+$ RNAs are synthesized by in vitro transcription from non-coding yeast genomic DNA. After incubation at 37° C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85° C to the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is

then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μl 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 μg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in U.S. Patent No. 5,807,522, incorporated herein by reference. 1 μl of the array element DNA, at an average concentration of 100 ng/μl, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60° C followed by washes in 0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 μl of sample mixture consisting of 0.2 μg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65° C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm^2 coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 μl of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60° C. The arrays are washed for 10 min at 45° C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45° C in a second wash buffer (0.1X SSC), and dried.

Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

XI. Complementary Polynucleotides

Sequences complementary to the SECP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring SECP. For example, SEQ ID NO:135 and SEQ ID NO:136 are complementary polynucleotides to SEQ ID NO:114 and SEQ ID NO:116, respectively. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of SECP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the SECP-encoding transcript.

XII. Expression of SECP

Expression and purification of SECP is achieved using bacterial or virus-based expression systems. For expression of SECP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express SECP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of SECP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding SECP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, SECP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from

SECP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified SECP obtained by these methods can be used directly in the assays shown in Examples XVI, XVII, and XVIII, where applicable.

XIII. Functional Assays

SECP function is assessed by expressing the sequences encoding SECP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of SECP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding SECP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding SECP and other genes of interest can be analyzed by northern

analysis or microarray techniques.

XIV. Production of SECP Specific Antibodies

SECP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the SECP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-SECP activity by, for example, binding the peptide or SECP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XV. Purification of Naturally Occurring SECP Using Specific Antibodies

Naturally occurring or recombinant SECP is substantially purified by immunoaffinity chromatography using antibodies specific for SECP. An immunoaffinity column is constructed by covalently coupling anti-SECP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing SECP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of SECP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/SECP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and SECP is collected.

XVI. Identification of Molecules Which Interact with SECP

SECP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled SECP, washed, and any wells with labeled SECP complex are assayed. Data obtained using different concentrations of SECP are used to calculate values for the number, affinity, and association of SECP with the

candidate molecules.

Alternatively, molecules interacting with SECP are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989) Nature 340:245-246, or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

SECP may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

XVII. Demonstration of SECP Activity

An assay for growth stimulating or inhibiting activity of SECP measures the amount of DNA synthesis in Swiss mouse 3T3 cells (McKay, I. and Leigh, I., eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York, NY). In this assay, varying amounts of SECP are added to quiescent 3T3 cultured cells in the presence of [^{3}H]thymidine, a radioactive DNA precursor. SECP for this assay can be obtained by recombinant means or from biochemical preparations. Incorporation of [^{3}H]thymidine into acid-precipitable DNA is measured over an appropriate time interval, and the amount incorporated is directly proportional to the amount of newly synthesized DNA. A linear dose-response curve over at least a hundred-fold SECP concentration range is indicative of growth modulating activity. One unit of activity per milliliter is defined as the concentration of SECP producing a 50% response level, where 100% represents maximal incorporation of [^{3}H]thymidine into acid-precipitable DNA .

Alternatively, an assay for SECP activity measures the stimulation or inhibition of neurotransmission in cultured cells. Cultured CHO fibroblasts are exposed to SECP. Following endocytic uptake of SECP, the cells are washed with fresh culture medium, and a whole cell voltage-clamped Xenopus myocyte is manipulated into contact with one of the fibroblasts in SECP-free medium. Membrane currents are recorded from the myocyte. Increased or decreased current relative to control values are indicative of neuromodulatory effects of SECP (Morimoto, T. et al. (1995) Neuron 15:689-696).

Alternatively, an assay for SECP activity measures the amount of SECP in secretory, membrane-bound organelles. Transfected cells as described above are harvested and lysed. The lysate is fractionated using methods known to those of skill in the art, for example, sucrose gradient ultracentrifugation. Such methods allow the isolation of subcellular components such as the Golgi apparatus, ER, small membrane-bound vesicles, and other secretory organelles. Immunoprecipitations from fractionated and total cell lysates are performed using SECP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The concentration of SECP in secretory organelles relative to SECP in total cell lysate is

proportional to the amount of SECP in transit through the secretory pathway.

Alternatively, AMP binding activity is measured by combining SECP with ^{32}P-labeled AMP. The reaction is incubated at 37°C and terminated by addition of trichloroacetic acid. The acid extract is neutralized and subjected to gel electrophoresis to remove unbound label. The radioactivity retained in the gel is proportional to SECP activity.

Alternatively, SECP activity for SEQ ID NO:67, for example, can be measured as protease inhibitory activity. Trypsin (100 units) is incubated at ambient temperature in a quartz cuvette in pH 7.6 assay buffer containing 63 mM sodium phosphate, 0.23 mM N α-benzoyle-L-arginine ethyl ester, 0.06 mM hydrochloric acid, with or without SECP. Immediately after mixing by inversion, the increase in $A_{253\ nm}$ is recorded for approximately 5 minutes and the enzyme activity is calculated (Bergmeyer, H.U. et al. (1974) Meth. Enzym. Anal. 1:515-516). SECP activity is proportional to its effect on the activity of trypsin.

XVIII. Demonstration of Immunoglobulin Activity

An assay for SECP activity measures the ability of SECP to recognize and precipitate antigens from serum. This activity can be measured by the quantitative precipitin reaction. (Golub, E. S. et al. (1987) Immunology: A Synthesis, Sinauer Associates, Sunderland, MA, pages 113-115.) SECP is isotopically labeled using methods known in the art. Various serum concentrations are added to constant amounts of labeled SECP. SECP-antigen complexes precipitate out of solution and are collected by centrifugation. The amount of precipitable SECP-antigen complex is proportional to the amount of radioisotope detected in the precipitate. The amount of precipitable SECP-antigen complex is plotted against the serum concentration. For various serum concentrations, a characteristic precipitin curve is obtained, in which the amount of precipitable SECP-antigen complex initially increases proportionately with increasing serum concentration, peaks at the equivalence point, and then decreases proportionately with further increases in serum concentration. Thus, the amount of precipitable SECP-antigen complex is a measure of SECP activity which is characterized by sensitivity to both limiting and excess quantities of antigen.

Alternatively, an assay for SECP activity measures the expression of SECP on the cell surface. cDNA encoding SECP is transfected into a non-leukocytic cell line. Cell surface proteins are labeled with biotin (de la Fuente, M.A. et.al. (1997) Blood 90:2398-2405). Immunoprecipitations are performed using SECP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of SECP expressed on the cell surface.

Alternatively, an assay for SECP activity measures the amount of cell aggregation induced by overexpression of SECP. In this assay, cultured cells such as NIH3T3 are transfected with cDNA encoding SECP contained within a suitable mammalian expression vector under control of a strong

promoter. Cotransfection with cDNA encoding a fluorescent marker protein, such as Green Fluorescent Protein (CLONTECH), is useful for identifying stable transfectants. The amount of cell agglutination, or clumping, associated with transfected cells is compared with that associated with untransfected cells. The amount of cell agglutination is a direct measure of SECP activity.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
3211795	1	3211795CD1	68	3211795CB1
6813464	2	6813464CD1	69	6813464CB1
2156540	3	2156540CD1	70	2156540CB1
894939	4	894939CD1	71	894939CB1
4620890	5	4620890CD1	72	4620890CB1
5514146	6	5514146CD1	73	5514146CB1
7474769	7	7474769CD1	74	7474769CB1
065296	8	065296CD1	75	065296CB1
231994	9	231994CD1	76	231994CB1
538054	10	538054CD1	77	538054CB1
1259305	11	1259305CD1	78	1259305CB1
1483702	12	1483702CD1	79	1483702CB1
1519324	13	1519324CD1	80	1519324CB1
1630169	14	1630169CD1	81	1630169CB1
1664253	15	1664253CD1	82	1664253CB1
1864715	16	1864715CD1	83	1864715CB1
1929395	17	1929395CD1	84	1929395CB1
1987737	18	1987737CD1	85	1987737CB1
2122866	19	2122866CD1	86	2122866CB1
2123981	20	2123981CD1	87	2123981CB1
2200177	21	2200177CD1	88	2200177CB1
2319255	22	2319255CD1	89	2319255CB1
2792452	23	2792452CD1	90	2792452CB1
2853088	24	2853088CD1	91	2853088CB1
2949004	25	2949004CD1	92	2949004CB1
3011670	26	3011670CD1	93	3011670CB1
3242083	27	3242083CD1	94	3242083CB1
3363391	28	3363391CD1	95	3363391CB1
3703614	29	3703614CD1	96	3703614CB1
4000975	30	4000975CD1	97	4000975CB1
4598831	31	4598831CD1	98	4598831CB1

Table 1 (cont.)

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
4992201	32	4992201CD1	99	4992201CB1
5441583	33	5441583CD1	100	5441583CB1
1639243	34	1639243CD1	101	1639243CB1
1335166	35	1335166CD1	102	1335166CB1
166894	36	166894CD1	103	166894CB1
217969	37	217969CD1	104	217969CB1
335237	38	335237CD1	105	335237CB1
938306	39	938306CD1	106	938306CB1
1448129	40	1448129CD1	107	1448129CB1
1761049	41	1761049CD1	108	1761049CB1
1959587	42	1959587CD1	109	1959587CB1
2303463	43	2303463CD1	110	2303463CB1
2512281	44	2512281CD1	111	2512281CB1
2755924	45	2755924CD1	112	2755924CB1
2796369	46	2796369CD1	113	2796369CB1
3010920	47	3010920CD1	114	3010920CB1
3360955	48	3360955CD1	115	3360955CB1
3409459	49	3409459CD1	116	3409459CB1
4102938	50	4102938CD1	117	4102938CB1
4124601	51	4124601CD1	118	4124601CB1
4180577	52	4180577CD1	119	4180577CB1
5265807	53	5265807CD1	120	5265807CB1
5405979	54	5405979CD1	121	5405979CB1
7481109	55	7481109CD1	122	7481109CB1
6247114	56	6247114CD1	123	6247114CB1
3243866	57	3243866CD1	124	3243866CB1
7475633	58	7475633CD1	125	7475633CB1
1431268	59	1431268CD1	126	1431268CB1
2414185	60	2414185CD1	127	2414185CB1
5266594	61	5266594CD1	128	5266594CB1
7610617	62	7610617CD1	129	7610617CB1

Table 1 (cont.)

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
1902436	63	1902436CD1	130	1902436CB1
2310369	64	2310369CD1	131	2310369CB1
6180576	65	6180576CD1	132	6180576CB1
2274523	66	2274523CD1	133	2274523CB1
1801820	67	1801820CD1	134	1801820CB1

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
1	3211795CD1	g5764665	3.80E-156	[Homo sapiens] cerebral cell adhesion molecule Starzyk, R.M. et al. (2000) J. Infect. Dis. 181:181-187
2	6813464CD1	g14794726	0	[fl][Homo sapiens] CUB and sushi multiple domains 1 protein Sun, P. C. et al. (2001) Transcript map of the 8p23 putative tumor suppressor region. Genomics. 75:17-25
3	2156540CD1	g7259265	2.40E-28	[Mus musculus] contains transmembrane (TM) region Inoue, S. et al. (2000) Growth suppression of Escherichia coli by induction of expression of mammalian genes with transmembrane or ATPase domains. Biochem. Biophys. Res. Commun. 268:553-561
4	894939CD1	g7242876	4.30E-171	[Mus musculus] kidney predominant protein
5	4620890CD1	g13874437	1.00E-134	[fl][Homo sapiens] cerebral protein-11
6	5514146CD1	g1256001	5.70E-46	[Homo sapiens] LIV-1 protein
7	7474769CD1	g13603845	0	[fl][Mus musculus] ribonuclease/angiogenin inhibitor 2 Wang, P. J. et al. (2001) An abundance of X-linked genes expressed in spermatogonia. Nat. Genet. 27:422-426
34	1639243CD1	g927209	5.00E-20	[Homo sapiens] alpha 1C adrenergic receptor isoform 2
35	1335166CD1	g3002527	1.80E-21	[Homo sapiens] neuronal thread protein AD7c-NTP
57	3243866CD1	g6273399	3.30E-27	[Homo sapiens] melanoma-associated antigen MG50 Weiler, S. R. et al. (1994) Assignment of a human melanoma associated gene MG50 (D2S448) to chromosome 2p25.3 by fluorescence in situ hybridization. Genomics 22:243-244
58	7475633CD1	g3097285	1.20E-65	[Rattus norvegicus] ZOG

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	3211795CD1	622	S125 S138 S153 S262 S272 S383 S542 S569 S591 S612 T300 T406 T562 T584 Y126 Y140 Y281 Y492	N184 N381 N96	Signal_cleavage: M1-G29	SPSCAN
					Signal peptide: M1-P32	HMMER
					Lysyl hydrolase Lysyl_hydro: V165-I240	HMMER_PFAM
					PROCOLLAGENLYSINE 2OXOGLUTARATE 5DIOXYGENASE PRECURSOR LYSYL HYDROXYLASE OXIDOREDUCTASE DIOXYGENASE SIGNAL IRON PD009947: P52-L291	BLAST_PRODOM
					LYSYL HYDROXYLASE CHAIN DM07920|Q02809|1-726: R155-L291	BLAST_DOMO
2	6813464CD1	529	S109 S213 S243 S251 S299 S56 T113 T192 T273 T405 T416 T99	N312 N411	Signal peptide: M1-A21	SPSCAN
					Signal peptide: M1-S22	HMMER
					CUB domains: C136-Y244, C309-F414	HMMER_PFAM
					Sushi domain (SCR repeat): C74-C131, C252-C305, C422-C479	HMMER_PFAM
					Sushi domain proteins PF00084: G93-Y104, N296-C305	BLIMPS_PFAM
					GLYCOPROTEIN DOMAIN EGFLIKE PROTEIN PRECURSOR SIGNAL RECEPTOR INTRINSIC FACTORB12 REPEAT PD000165:S144-Y244, T314-F414	BLAST_PRODOM
					C1R/C1S REPEAT DM00162|I49540|748-862: C136-Y244, C309-F410 DM00162|P98063|755-862: S144-Y244, T314-F410 DM00162|I49540|592-708:C305-F414 DM00162|P13497|595-703:T314-F414	BLAST_DOMO

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
3	2156540CD1	204	S101 S124 S196 S63 S83 T46 T51 T54 T88 Y140	N122	signal_peptide: M1-A24	HMMER
					signal_cleavage: M1-P27	SPSCAN
					Immunoglobulin domain ig: G55-V144	HMMER_PFAM
4	894939CD1	406	S151 S200 S226 S249 S255 S288 S400	N134 N159 N187 N230 N333 N65 N95	signal_peptide: M1-G35	HMMER
					signal_cleavage: M1-G35	SPSCAN
					transmembrane_domain: L369-G387	HMMER
					Leucine_Zipper: L371-L392	HMMER
5	4620890CD1	477	S151 S166 S183 S226 S235 S25 S253 S279 S360 S388 S70 T13 T325 T87	N201 N270	signal_peptide: M439-C460	HMMER
					transmembrane_domain: I418-F435	MOTIFS
					PROTEIN TESTISSPECIFIC TEX28KIAA0481 HH1480 PD156308: P55-Q145 PD184097: L287-W463	BLAST_PRODOM
6	5514146CD1	691	S130 S136 S149 S160 S250 S293 S30 S312 S329 S392 S470 S51 S513 S530 S565 S598 S645 T24 T246 T26 T362 T658 T67	N134 N162 N201 N242 N290 N59	signal_cleavage: M1-S23	SPSCAN
					transmembrane_domain: I603-L621, W632-P651, P661-Y685	HMMER
					PROTEIN GUFA TRANSMEMBRANE MEMBRANE INTERGENIC REGION INNER CONSERVED SIMILARITY MYXOCOCCUS PD004603: K533-L682	BLAST_PRODOM

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
7	7474769CD1	919	S125 S146 S157 S230 S275 S37 S393 S40 S462 S523 S543 S591 S632 S638 S661 S733 T217 T574 T679 T751 T814 T885 T896 T906 T915 Y109 Y47 Y645	N228 N601 N682 N731 N852 N904	signal_cleavage: M1-T16	SPSCAN
					RIBONUCLEASE INHIBITOR REPEAT LEUCINEREPEAT 3DSTRUCTURE PLACENTAL RAI RI RECEPTOR RIBONUCLEASE/ANGIOGENIN PD017636: L737-L842, L794-E889, L681-S787	BLAST_PRODOM
					Atp_Gtp_A: G80-T87	MOTIFS
8	065296CD1	178	S24 S7 S79 S81 S87	N116 N74 N75	signal_cleavage: M1-G26	SPSCAN
9	231994CD1	310	S281 S93 T136	N101 N219	signal_cleavage: M1-P24	SPSCAN
10	538054CD1	559	S121 S150 S216 S258 S294 S385 S474 S71 T274 T28 T363 T41	N237 N26 N292 N346	signal_peptide: M1-V24	HMMER
					signal_cleavage: M1-N26	SPSCAN
					transmembrane_domain: M1-I25	HMMER
					BRUSH BORDER 61.9KD PROTEINPRECURSOR SIGNAL PD144534: L8-C555	BLAST_PRODOM
11	1259305CD1	477	S113 S129 S211 S273 S471 T135 T139 T142 T18 T338	N111 N209	signal_cleavage: M1-S40	SPSCAN
					transmembrane_domain: M23-A39, F56-I74, F250-L270, M315-L335, W406-N430	HMMER
12	1483702CD1	176	S108 S132 S138 S142 S3 S34 T32	N47	signal_peptide: M111-S132	HMMER
13	1519324CD1	190	S119 S164 S189 S23 S86 S93 S98		signal_cleavage: M1-S23	SPSCAN
14	1630169CD1	75		N50	signal_cleavage: M1-A27	SPSCAN
					signal_peptide: M1-S22	HMMER
					transmembrane_domain: L3-S22	HMMER
					sulfatases signature sulfatase_2.prf: S22-W76	PROFILESCAN

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
15	1664253CD1	265	S12 S14 S187 S30 S64 T162	N197	signal_peptide: M116-A264	HMMER
16	1864715CD1	202	S33 T167 T65	N10	signal_peptide: M1-A19	HMMER
					signal_cleavage: M1-G15	SPSCAN
					transmembrane_domain: I173-R193	HMMER
17	1929395CD1	111	T4 T67		signal_cleavage: M1-G34	SPSCAN
18	1987737CD1	105	S75 T38		signal_cleavage: M1-G28	SPSCAN
19	2122866CD1	717	S172 S3 S359 S397 S454 S46 S469 S488 S5 S526 S567 S594 S7 T117 T20 T410 Y213	N130 N274 N44 N520 N666	signal_cleavage: M133-P197	SPSCAN
					PROTEIN FACTOR REPEAT CELL CHROMOSOME INTERGENIC REGION HOST C1 JASMONATE PD013240: N149-N306	BLAST_PRODOM
					RGD: R432-D434	MOTIFS
20	2123981CD1	580	S109 S120 S210 S260 S322 T121 T178 T272 T31 T327 T360 T403 T415 T450 T460 T545 T70 T73 T75 Y522	N276 N337 N543 N99	signal_peptide: M1-K21	HMMER
					signal_cleavage: M1-A16	SPSCAN
21	2200177CD1	172	S113 S148 S93 T138 Y121		signal_cleavage: M1-C50	SPSCAN
					transmembrane_domain: L34-I51	HMMER
22	2319255CD1	256	S242 S49 T116 T130 T3 T66	N200	signal_cleavage: M1-A30	SPSCAN
23	2792452CD1	93	S15 S24		signal_peptide: M1-S22	HMMER
					transmembrane_domain: L41-I58	HMMER
					signal_cleavage: M1-G52	SPSCAN
24	2853088CD1	112	S89		signal_peptide: M1-G36	HMMER
					signal_cleavage: M1-A27	SPSCAN
					J SP NEURONAL THREAD PD003801: P30-R71	BLAST_PRODOM

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
25	2949004CD1	186	S26 S91 S94 T19 T65		signal_peptide: M75-S94	HMMER
					transmembrane_domain: I139-L163	HMMER
					signal_peptide: M128-A150	SPSCAN
26	3011670CD1	487	S356 S36 S445 S89 T227 T262 T313 T44 T60 T91 Y66	N135 N203 N34 N397	transmembrane_domain: W106-N126, I136-L155, V170-V188, M306-C323, L459-F479	HMMER
					Atp_Gtp_A: G53-T60	MOTIFS
					signal_cleavage: M1-A66	SPSCAN
27	3242083CD1	350	S311 T261 T73	N294	transmembrane_domain: V43-W65, M84-C102, F172-V191	HMMER
					signal_peptide: M144-R162	HMMER
28	3363391CD1	450	S132 S140 S196 S247 S289 S380 S385 S417 S428 T208 T214 T357 T445 Y368 Y403	N131 N206 N287	transmembrane_domain: L146-R162	HMMER
					signal_peptide: M173-T193	HMMER
29	3703614CD1	400	S372 S71 S82 T32 T348 T358 Y362 Y76		transmembrane_domain: M173-Y190, P198-V221, L293-V310	HMMER
					signal_peptide: M1-A24	SPSCAN
30	4000975CD1	133	S54 S7		signal_cleavage: M1-M25	HMMER
					signal_peptide: M327-A344	BLAST_DOMO
31	4598831CD1	359	S139 S300 S359 T355 T41 T63		do MEMBRANE; YOL002C; CHROMOSOME; C30D11.11; DM02642|Q09749|49-323: Y55-G321 DM02642|S62569|169-441: T41-H291 DM02642|Q09910|169-441: T41-H291 DM02642|S61982|50-325: D46-Q310	
					signal_cleavage: M1-G65	SPSCAN
32	4992201CD1	72	S17 S32 T7	N14	transmembrane_domain: M39-F58	HMMER
					signal_peptide: M53-A80	HMMER
33	5441583CD1	112			Heme oxygenase signature heme_oxygenase.prf: T9-M53	PROFILESCAN

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
34	1639243CD1	149	S93		Signal cleavage: M1-Q45	SPSCAN
					J SP NEURONAL THREAD PD003801:R89-I130	BLAST_PRODOM
35	1335166CD1	97	S30 S36 S52 S70		Signal peptide: M1-P21	HMMER
					Signal cleavage: M1-A43	SPSCAN
					RNA EDITING PROTOONCOGENE REPEAT PD005171:S3-P45	BLAST_PRODOM
36	166894CD1	104	S25 T60		Signal cleavage: M1-T27	SPSCAN
					Signal peptide: M1-S29	HMMER
					Insulin-like growth factor binding proteins IGFB: C80-M93	HMMER_PFAM
					P-II protein signatures pii_glnb_cter.prf: L20-G76	PROFILESCAN
					Ribosomal protein S19e signature ribosomal_s19e.prf: S24-P77	PROFILESCAN
					Plant PEC family metallothionein signature PR00877 L75-R82	BLIMPS_PRINTS
37	217969CD1	99	T9		Signal cleavage: M1-S35	SPSCAN
					RNA EDITING PROTOONCOGENE REPEAT PD005171: F29-L66 P_value 2.6e-07	BLAST_PRODOM
38	335237CD1	80	S38	N42	Signal peptide: M1-S18	HMMER
					Signal cleavage: M1-S18	SPSCAN
39	938306CD1	96	S41 S45 T79	N58	Signal peptide: M1-G18	HMMER
					Signal cleavage: M1-G18	SPSCAN
40	1448129CD1	92	S62		Signal peptide: M1-A28	HMMER
					Signal cleavage: M1-G31	SPSCAN
					Aminotransferases class-I pyridoxal-phosphate attachment site asp_aminotransferase.prf: K55-A92	PROFILESCAN
					Respiratory-chain NADH dehydrogenase 75 Kd subunit signatures complex1_75k_3.prf: F5-Q79	PROFILESCAN

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
41	1761049CD1	77	S37		Signal cleavage: M1-C66	SPSCAN
					Transmembrane domain: Y40-M64	HMMER
42	1959587CD1	75	S48 S70		Signal cleavage: M1-G26	SPSCAN
					Signal peptide: M1-G23	HMMER
43	2303463CD1	85			Signal cleavage: M1-A30	SPSCAN
					Transmembrane domain: L22-A48	HMMER
44	2512281CD1	89	S61 S62 T29 T42 T58		Signal peptide: M1-R31	HMMER
					Signal cleavage: M1-P22	SPSCAN
					5-hydroxytryptamine 1B receptor PR00513 R31-T42	BLIMPS_PRINTS
45	2755924CD1	123			Signal peptide: M1-A19	HMMER
					Signal cleavage: M1-S24	SPSCAN
					Transmembrane domain: I38-A53	HMMER
46	2796369CD1	159	S97 T35		Signal cleavage: M1-G27	SPSCAN
					Tyrosine specific protein phosphatases active site tyr_phosphatase.prf: L42-E100	PROFILESCAN
47	3010920CD1	77	S7		Signal cleavage: M1-T33	SPSCAN
48	3360955CD1	130	S15 S32 S95 T25		Signal cleavage: M1-T14	SPSCAN
					Orexin receptor signature PR01064 N78-T89	BLIMPS_PRINTS
49	3409459CD1	97			Signal peptide: M1-G20	HMMER
					Signal cleavage: M1-G20	SPSCAN
50	4102938CD1	74	S3		Signal cleavage: M1-L25	SPSCAN
					Signal peptide: M1-G33	HMMER
					Hemerythrins signature hemerythrins.prf: Q21-L60	PROFILESCAN
51	4124601CD1	74	S12		Signal cleavage: M1-G46	SPSCAN
					Annexin Type III Signature PR00199 W17-A26	BLIMPS_PRINTS

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
52	4180577CD1	151	S115 S116 S128 S63 S68 S77 T93	N40	Signal peptide: M1-A22	HMMER
					Signal cleavage: M1-S21	SPSCAN
53	5265807CD1	137	S123 S79		Signal cleavage: M1-A42	SPSCAN
					Signal peptide: M1-G31	HMMER
					Ank repeat ank: Y13-K45 L46-R78 S79-Y113	HMMER_PFAM
					Ank repeat PD0007 L77-A89	BLIMPS_PRODOM
					Ank repeat protein motif PF00023 G80-H89	BLIMPS_PFAM
54	5405979CD1	137	S48 T118		Signal peptide: M1-A24	HMMER
					Signal cleavage: M1-L16	SPSCAN
55	7481109CD1	205	T181 T31 Y141	N131 N143	Signal peptide: M2-Q22	HMMER
					Signal cleavage: M1-F26	SPSCAN
					HYPERGLYCEMIC HORMONE TY PR00548: L158-C168	BLIMPS_PRINTS
					Pancreatic ribonucleases RNaseA: F104-C168	HMMER_PFAM

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
56	6247114CD1	199	S108 S113 S123 S136 S146 S42 S67 S91 T162 T175 T22 T31 T77	N111 N61 N89	Signal peptide: M1-S21 Signal cleavage: M1-A18 PANCREATIC RIBONUCLEASE FAMILY DM00621|P39873|28-166: S73-V183 Pancreatic ribonuclease family signature rnase_pancreatic.prf: G92-K138	HMMER SPSCAN BLAST_DOMO PROFILESCAN
57	3243866CD1	719	S108 S204 S255 S276 S400 S410 S618 S640 S685 S692 S76 T141 T198 T270 T305 T348 T454 T467 T480 T510 T647 T661 T668 T691 T704 T713	N330 N339 N382 N406 N452 N671 N672 N73	signal_peptide: M1-P21 signal_cleavage: M1-A17 transmembrane_domain: M529-V554 Polycystic kidney disease protein signature PR00500B: P241-E261 Leucine-rich repeat signature PR00019A: L149-I162 Leucine Rich Repeat LR: S172-H195, K196-A219, R52-T75, S76-R99, N100-S123, N124-F147, A148-V171 Leucine rich repeat C-terminal domain LRRCT: N240-E285 Immunoglobulin domain ig: G301-A359 RGD: R311-D313	HMMER SPSCAN HMMER BLIMPS_PRINTS BLIMPS_PRINTS HMMER_PFAM HMMER_PFAM HMMER_PFAM MOTIFS

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
58	7475633CD1	383	S290 S31 S45 T159 T188 T258 T328	N157	signal_peptide: M1-A26	HMMER
					signal_cleavage: M1-G21	SPSCAN
					transmembrane _domain: V308-R330	HMMER
					EGF DM00003|I48324|235-285: C40-K90 DM00003|Q07645|107-174: G113-V173 DM00003|P10041|275-331: W75-K131 DM00003|P80370|72-127: Q76-K131	BLAST_DOMO
					Laminin-type EGF-like domain BL01248 C48-C60	BLIMPS_BLOCKS
					Type II EGF-like signature PR00010C: N194-F204	BLIMPS_PRINTS
					Type III EGF-like signature PR00011D: C39-C57	BLIMPS_PRINTS
					EGF-like domain EGF: C216-C247, C29-C57, C60-C88, C95-C128, C135-C171, C178-C209	HMMER_PFAM
					EGF-like domain Egf: C46-C57, C77-C88, C117-C128, C160-C171, C198-C209, C236-C247	MOTIFS
					Zinc_Finger_C3hc4: C7-I16	MOTIFS
					Asp or Asn hydroxylation site found in EGF-domain proteins Asx_Hydroxyl: C189-C200, C227-C238	MOTIFS
59	1431268CD1	126	S2 S54		signal_peptide: M1-A35	HMMER
60	2414185CD1	137	T65		signal_peptide: M1-G21	HMMER
61	5266594CD1	77	T35		signal_peptide: M1-G29	HMMER
62	7610617CD1	110	S46 T83		signal_peptide: M1-S21	HMMER

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
63	1902436CD1	103	T72		signal_peptide: M1-S18	HMMER
					signal_cleavage: M1-S18	SPSCAN
					transmembrane_domain: I4-I23	HMMER
64	2310369CD1	192	S153 S28 S56 T186 T57		signal_cleavage: M1-A32	SPSCAN
					transmembrane_domain: V114-T134	HMMER
65	6180576CD1	310	S154 S158 S201 S307 S5 S79 S93 T225 T253 T55 T71 Y163		signal_cleavage: M1-G61	SPSCAN
66	2274523CD1	135	S12		signal_cleavage: M1-S44	SPSCAN
67	1801820CD1	205	S110 S126 S198 S68 T185 T34	N67	signal_cleavage: M1-K36	SPSCAN
					Bowman-Birk serine protease inhibitor family signature bowman_birk.prf: C104-C194	PROFILESCAN

Table 4

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
68	3211795CB1	2569	1-513, 2404-2569	8052177J1 (FTUBTUE01)	1775	2401
				6756010H1 (SINTFER02)	453	1192
				6756010J1 (SINTFER02)	1047	1810
				6888428H1 (BRAITDR03)	1419	1955
				6944985H1 (FTUBTUR01)	1950	2416
				2185379F6 (PROSNOT26)	2103	2569
				6900869H1 (MUSLTDR02)	729	1401
				7317029H2 (BONFTXT01)	1	560
69	6813464CB1	2387	104-184, 546-2387	GNN.g8051897_000002_002	185	519
				GNN.g9408716_002.edit	1	397
				71926854V1	1924	2387
				8085111H1 (BRACDIK08)	355	919
				71926415V1	1527	2107
				71930431V1	1025	1536
				71928622V1	1515	2047
				6813464F6 (ADRETUR01)	559	1255
70	2156540CB1	1959	1918-1959, 723-1073	g2204647	1385	1959
				71147007V1	1221	1916
				7714520H1 (SINTFEE02)	705	1255
				g5553166	1	422
				6269361H1 (BRAIFEN03)	1255	1927
				7192294H2 (BRATDIC01)	31	462
				7460559H1 (LIVRTUE01)	264	948
				825696H1 (PROSNOT06)	1674	1941
71	894939CB1	1562	1-253	71113024V1	1	569
				71113961V1	935	1562
				71111802V1	582	1258
				71178768V1	500	1205

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
72	4620890CB1	3425	1-355, 1127-2529	3730714F6 (SMCCNON03)	3124	3425
				7126752H1 (COLNDIY01)	2199	2753
				71260484V1	964	1582
				4041777H1 (BRAITUT26)	1	299
				7089611R8 (BRAUTDR03)	439	1129
				7186380H1 (BONRFEC01)	1573	2194
				5608905T6 (MONOTXS05)	1116	1696
				6216030H1 (MUSCDIT06)	2803	3267
				7089611F8 (BRAUTDR03)	56	726
				5721967H1 (SEMVNOT05)	2457	2825
				6365636F6 (ARTANOT07)	1758	2350
				4058011F6 (SPLNNOT13)	2882	3293
73	5514146CB1	3130	2654-3130, 1611-2305, 992-1198	70403184D1	2245	2786
				70401863D1	2527	3130
				70942023V1	1409	2085
				6770687H1 (BRAUNOR01)	239	723
				7998254H1 (BRAITUC02)	1	446
				5514146F6 (BRADDIR01)	1172	1710
				70941524V1	1731	2281
				4253735F6 (BRADDIR01)	564	1091
				5511496F8 (BRADDIR01)	813	1434
74	7474769CB1	3172	646-970, 2450-2614, 1939-2217	55077715J1	1	392
				5544880H1 (TESTNOC01)	1456	1649
				g2100308	2911	3172
				GNN.g6693125_000016_002	200	2789
75	065296CB1	2094	1826-2094	8038136J1 (SMCRUNE01)	1273	2094
				71058917V1	872	1524
				g4393469	1	357
				71247541V1	846	1456
				71059505V1	219	883

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
76	231994CB1	1119	1-163	6424977H1 (BRSTUNT01)	19	703
				1735672F6 (COLNNOT22)	309	846
				3484368H1 (KIDNNOT31)	1	154
				404248R6 (TMLR3DT01)	620	1119
77	538054CB1	3321	1513-1629, 2956-3321	2304848R6 (NGANNOT01)	1671	2208
				70558993V1	546	1031
				70561551V1	403	982
				6951008H1 (BRAITDR02)	1039	1798
				530151R6 (BRAINOT03)	2887	3321
				6884875H1 (BRAHTDR03)	984	1466
				7262670H1 (UTRETMC01)	1	433
				4880167H1 (UTRMTMT01)	2499	2782
				351423T6 (LVENNOT01)	2563	3147
				6995967H1 (BRAXTDR17)	1820	2548
78	1259305CB1	2646	1-1540	661278T6 (BRAINOT03)	1959	2624
				6782356J1 (OVARDIR01)	584	1338
				8009718H1 (NOSEDIC02)	1	556
				71294610V1	2104	2646
				6782356H1 (OVARDIR01)	1332	2025
				70979178V1	1068	1745
				70978234V1	349	977
79	1483702CB1	1749	1712-1749, 1130-1181	SXAF03884V1	1	513
				1478338F1 (CORPNOT02)	404	960
				1477349F2 (CORPNOT02)	1178	1731
				1832249R6 (BRAINON01)	1421	1749
				2571527R6 (HIPOAZT01)	713	1222
80	1519324CB1	2339	1987-2339, 1-66	569251H1 (MMLR3DT01)	1203	1522
				4435164H2 (LUNGNOT38)	918	1216
				268264H1 (HNT2NOT01)	473	759
				3504827H1 (ADRENOT11)	2031	2319
				3803618H1 (BLADTUT03)	1521	1831

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
				3188988H1 (THYMNON04)	1459	1783
				1556067T6 (BLADTUT04)	1662	2316
				1484363F6 (CORPNOT02)	1	478
				1347568H1 (PROSNOT11)	2175	2339
				2959710H2 (ADRENOT09)	379	692
				4103403F6 (BRSTTUT17)	597	1168
				3802959H1 (BLADTUT03)	1197	1499
81	1630169CB1	1006	1-1006	6771071J1 (BRAUNOR01)	1	547
				6915144J1 (PITUDIR01)	386	815
				71296019V1	654	1006
82	1664253CB1	1050	1-65, 979-1050	6975645F8 (BRAHTDR04)	571	1050
				1511079H1 (LUNGNOT14)	1	206
				7988873H1 (UTRSTUC01)	280	911
				6975645R8 (BRAHTDR04)	153	745
83	1864715CB1	1774	1-91, 864-1189	2866801H1 (KIDNNOT20)	505	837
				3736166H1 (SMCCNOS01)	1503	1774
				2972289H2 (HEAONOT02)	168	461
				1864715T6 (PROSNOT19)	1175	1760
				4585982H1 (OVARNOT13)	928	1195
				1864715F6 (PROSNOT19)	396	828
				g5431009	1	452
				6521196H1 (CONFTDT02)	584	1192
84	1929395CB1	2608	1-230, 2378-2608, 1364-1424, 1859-1908	6426607H1 (LUNGNON07)	223	830
				7159765H1 (HNT2TXC01)	1	437
				6403390H1 (UTRENOT10)	2330	2608
				5203330H2 (STOMNOT08)	1168	1429
				8184678H1 (EYERNON01)	1894	2382
				999772R1 (KIDNTUT01)	1416	1963
				5776203H1 (BRAINOT20)	806	1360
85	1987737CB1	1336	1-362, 396-455	2569101R6 (HIPOAZT01)	1	375
				1987737R6 (LUNGAST01)	855	1336

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
				7651140J1 (STOMTDE01)	287	883
				70822631V1	379	1013
86	2122866CB1	3062	2470-3062	1344981T6 (PROSNOT11)	2525	3062
				3373914T6 (CONNTUT05)	1720	2315
				5106924H1 (PROSTUS19)	1482	1736
				70581466V1	571	1299
				7760472J1 (THYMNOE02)	1	697
				1861920T6 (PROSNOT19)	1986	2566
				888838R1 (STOMTUT01)	1125	1703
87	2123981CB1	2543	2458-2543	2502591T6 (CONUTUT01)	1788	2498
				70818388V1	760	1446
				2123981F6 (BRSTNOT07)	189	695
				7184911H1 (BONRFEC01)	1	201
				7687944J1 (PROSTME06)	658	1336
				7182169H1 (BONRFEC01)	104	694
				1746464F6 (STOMTUT02)	1479	2021
				1782982R6 (PGANNON02)	2215	2543
				70818385V1	1272	1844
				1570015F6 (UTRSNOT05)	2016	2528
88	2200177CB1	873	1-873	70715536V1	229	873
				2200177T6 (SPLNFET02)	1	653
89	2319255CB1	2134	1-1366	70972802V1	747	1350
				70972902V1	1606	2134
				2319255R6 (OVARNOT02)	1205	1770
				71291442V1	1388	1893
				70973293V1	369	927
				2719008F6 (THYRNOT09)	1	541
90	2792452CB1	1886	1-675, 1174-1420, 1824-1886	2631710F6 (COLNTUT15)	1358	1886
				2792452T6 (COLNTUT16)	1154	1816
				7056068H1 (BRALNON02)	1	611
				70747733V1	764	1290

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
				70750993V1	541	1199
91	2853088CB1	928	1-368, 541-928	6909603F6 (PITUDIR01)	403	928
				2853088F6 (BRSTTUT13)	1	510
92	2949004CB1	962	1-73, 524-574	6406737H1 (UTREDIT10)	207	755
				5773379H1 (BRAINOT20)	103	676
				5571763H1 (TLYMNOT08)	1	212
				5763826H1 (PROSBPT02)	569	962
93	3011670CB1	2644	2581-2644, 2402-2433, 269-296, 1289-1370	1454503T6 (PENITUT01)	1958	2571
				70529802V1	492	1220
				70527923V1	1101	1684
				70529939V1	1266	2004
				6784388H1 (SINITMC01)	270	744
				2130537H1 (KIDNNOT05)	2381	2644
				2493973H1 (ADRETUT05)	1	317
94	3242083CB1	1875	1605-1624, 760-1228, 140-182	8035143H1 (SMCRUNE01)	641	1410
				70729684V1	1272	1875
				7764378H1 (URETTUE01)	549	1215
				8035143J1 (SMCRUNE01)	1	640
95	3363391CB1	2378	1-1315	7407517H1 (UTREDME05)	1294	1790
				1691309F6 (PROSTUT10)	1890	2378
				4647910F6 (PROSTUT20)	1	702
				70701284V1	1462	1803
				7600868H1 (ESOGTME01)	415	935
				7600868J1 (ESOGTME01)	701	1335
				2055126T6 (BEPINOT01)	1750	2348
96	3703614CB1	1597	1-230, 1513-1597	70977847V1	1	619
				70979848V1	615	1114
				70977881V1	579	1096
				70977648V1	1047	1597
97	4000975CB1	653	1-653	70255942V1	1	528
				70256236V1	62	653

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
98	4598831CB1	3090	1-265, 1019-1311, 3031-3090	6450571H1 (BRAINOC01)	2461	3039
				7604731H1 (COLRTUE01)	277	871
				7280444H1 (BMARTXE01)	1	363
				7604731J1 (COLRTUE01)	1034	1618
				6531268H1 (LUNPTMC01)	2134	2785
				1717922F6 (UCMCNOT02)	1652	2072
				g884719	2646	3090
				7069481H1 (BRAUTDR02)	498	1097
				7078716H1 (BRAUTDR04)	1807	2285
				2598065F6 (OVARTUT02)	2715	3063
				7361768H1 (BRAIFEE05)	1093	1715
99	4992201CB1	1274	1-132, 1180-1274	7408255H1 (UTREDME05)	190	754
				6169092H1 (UTRSTDT01)	1004	1274
				264080H1 (HNT2AGT01)	1	357
				4992201F6 (LIVRTUT11)	356	1023
100	5441583CB1	1514	1-341, 1360-1514, 644-737	7619787J1 (KIDNTUE01)	1036	1514
				3216178F6 (TESTNOT07)	1	608
				71580026V1	543	1108
				71579908V1	651	1167
101	1639243CB1	1380	1332-1380, 1-505	2789856F6 (COLNTUT16)	1188	1375
				8215917H1 (FIBRTXC01)	435	1224
				71249544V1	1	632
				g1447886	892	1380
				1659932T6 (URETTUT01)	658	1319
102	1335166CB1	942	1-356, 806-942	6929148H1 (SINITMR01)	374	942
				1335166F6 (COLNNOT13)	1	512
103	166894CB1	1815	1-488, 1719-1815	2914860T6 (THYMFET03)	1046	1631
				8056981J1 (ESOGTUE01)	1	641
				6623672H1 (LIVRTMR01)	1310	1815
				5880645F8 (LIVRNON08)	516	1276

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
104	217969CB1	1120	801-1120, 1-542	1406479T6 (LATRTUT02)	238	897
				1406479F6 (LATRTUT02)	1	500
				7127039H1 (COLNDIY01)	546	1120
105	335237CB1	535	1-120, 389-535	6516316H1 (THYMDIT01)	1	517
				335237R6 (EOSIHET02)	188	535
106	938306CB1	1188	1-36, 806-1010, 232-330	6837739H1 (BRSTNON02)	580	1188
				1843293R6 (COLNNOT08)	1	610
107	1448129CB1	638	1-162	1448129F6 (PLACNOT02)	1	450
				g5152438	219	638
				1448129T6 (PLACNOT02)	217	599
108	1761049CB1	648	152-576, 1-91, 626-648	3335890F7 (BRAIFET01)	148	648
				55079223J1	1	647
109	1959587CB1	1181	1-484	2418384T6 (HNT3AZT01)	623	1181
				6562708H1 (MCLDTXT04)	227	911
				1959587H1 (BRSTNOT04)	1	228
				3087044F6 (HEAONOT03)	77	657
110	2303463CB1	1291	765-815, 965-1000, 1141-1291	70726533V1	587	1253
				g2052638	678	1291
				70725786V1	659	1270
				5371544F6 (BRAINOT22)	1	626
				4533095F8 (MYEPTXT01)	758	1287
111	2512281CB1	594	324-594	2512281F6 (CONUTUT01)	1	522
				7177623T8 (BRAXDIC01)	201	594
112	2755924CB1	852	1-173, 290-852	70252828V1	384	852
				71442317V1	1	603
113	2796369CB1	1361	1-244, 802-1135	5289268F6 (LIVRTUS02)	1	554
				5716622H1 (PANCNOT16)	772	1361
				2796369F6 (NPOLNOT01)	562	1128
				6264155H1 (MCLDTXN03)	483	950

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
135	3010920CB1.comp	1650	1-235, 1416-1650	71297582V1	1	646
				3010920F6 (MUSCNOT07)	805	1441
				70990273V1	1154	1650
				70991132V1	624	1290
115	3360955CB1	1845	898-1404, 162-191, 1-29, 1714-1845	70138150V1	993	1531
				70772246V1	485	1001
				3360955F6 (PROSBPT02)	1272	1845
				70137857V1	21	501
				g793031	1	360
136	3409459CB1.comp	1061	1-627, 435-1061	2096304H1 (BRAITUT02)	1	230
				4560677H1 (KERATXT01)	942	1061
				8049214H1 (LUNGTUS02)	628	946
				3409459F6 (PROSTUS08)	377	798
				7590996H1 (LIVRNOC07)	134	641
117	4102938CB1	1085	971-1085, 1-257, 684-783, 464-486	71665573V1	330	1085
				4102938F6 (BRSTTUT17)	1	507
118	4124601CB1	870	285-870	71157275V1	1	520
				4124601F6 (BRSTTUT26)	200	870
119	4180577CB1	3394	1-766, 3366-3394, 1487-1519	72081891D1	2015	2579
				72079822D1	2570	3394
				7370367H1 (ADREFEC01)	2278	2821
				72360174D1	1351	2147
				7371968H2 (BRAIFEE04)	413	992
				7675267J2 (NOSETUE01)	1	607
				72360060D1	1405	2231
				72082904D1	721	1412
120	5265807CB1	2343	56-519, 671-1627	71163549V1	560	1218
				71397639V1	1942	2343
				71396875V1	1547	2197

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
				71163327V1	1285	1845
				7142204H1 (LIVRDIT07)	1183	1750
				7098309H1 (BRACDIR02)	1	574
				71166280V1	540	1183
121	5405979CB1	751	718-751	g6711145	1	240
				2009872R6 (TESTNOT03)	303	751
				2009872T6 (TESTNOT03)	17	749
122	7481109CB1	618	1-618	GNN.g9588388_010	1	618
123	6247114CB1	979	1-384	2008559R6 (TESTNOT03)	538	979
				6247114F8 (TESTNOT17)	1	741
124	3243866CB1	3012	1-43, 2836-2885, 2307-2520, 1275-1779	8184163H1 (EYERNON01)	1215	1932
				6768945J1 (BRAUNOR01)	1	569
				70960957V1	2437	3012
				70869814V1	1551	2189
				7056045H1 (BRALNON02)	2131	2749
				3243866F6 (BRAINOT19)	735	1313
				70870169V1	2283	2835
				4783652F6 (BRATNOT03)	503	1074
125	7475633CB1	1600	1-72	6874650F6 (EPIMUNN04)	1	911
				72379524V1	900	1600
				6874650T6 (EPIMUNN04)	746	1587
126	1431268CB1	1001	335-547	5223435H1 (OVARDIT07)	650	922
				6747053H1 (BRAFNOT02)	286	909
				4139565H1 (BRSTTMT01)	709	1001
				70255212V1	1	559
127	2414185CB1	1424	1-46, 297-1424	7034594R6 (SINTFER03)	1	607
				6897648H1 (LIVRTMR01)	899	1424
				2414185F6 (HNT3AZT01)	570	1052
128	5266594CB1	1282	1-34, 990-1282	7764520H1 (URETTUE01)	730	1282
				5545864T8 (TESTNOC01)	1	829
				7373102F6 (BRAIFEE04)	29	632

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
				7736685J1 (BRAITUE01)	348	963
				3881415H1 (SPLNNOT11)	721	1008
129	7610617CB1	642	314-642	7610617F6 (KIDCTME01)	97	642
				7610617R6 (KIDCTME01)	1	609
130	1902436CB1	1326	845-1326, 740-776, 40-705	70882432V1	717	1326
				2216468F6 (SINTFET03)	1	480
				70880085V1	538	986
				70879063V1	412	953
131	2310369CB1	1486	139-395, 67-106	2310369R6 (NGANNOT01)	107	557
				3375613H1 (CONNTUT05)	1	265
				72069492V1	784	1486
				6438845H1 (BRAENOT02)	252	854
132	6180576CB1	1523	1-40	6116528T8 (SINITMT04)	762	1496
				7179642H1 (BRAXDIC01)	526	1151
				6146183H1 (BRANDIT03)	1185	1523
				6577826F8 (BRANDIT04)	1	571
133	2274523CB1	848	87-567	2274523T6 (PROSNON01)	378	848
				2083066F6 (UTRSNOT08)	87	583
				2068885H1 (PROSNOT26)	1	348
134	1801820CB1	2758	2484-2758	1315807F1 (BLADTUT02)	2434	2758
				2645114F6 (OVARTUT03)	1991	2502
				1798924F6 (COLNNOT27)	556	1212
				7079130H1 (BRAUTDR04)	1	600
				7409525H1 (BRAIFEJ02)	631	1253
				6479333H1 (PROSTMC01)	910	1424
				2303084R6 (BRSTNOT05)	1294	1816
				2926367F6 (TLYMNOT04)	1693	2229

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID	Representative Library
68	3211795CB1	OVARTUE01
69	6813464CB1	ADRETUR01
70	2156540CB1	BRAINOT09
71	894939CB1	BONSTUT01
72	4620890CB1	MONOTXS05
73	5514146CB1	BRADDIR01
74	7474769CB1	TESTNOC01
75	065296CB1	PLACNOB01
76	231994CB1	BRSTTUT01
77	538054CB1	LVENNOT01
78	1259305CB1	KIDNNOT19
79	1483702CB1	CORPNOT02
80	1519324CB1	KIDNNOT09
81	1630169CB1	BRAGNON02
82	1664253CB1	BRAHTDR04
83	1864715CB1	PROSNOT19
84	1929395CB1	COLNTUT03
85	1987737CB1	BRAINON01
86	2122866CB1	LIVRNON08
87	2123981CB1	PGANNON02
88	2200177CB1	KERANOT01
89	2319255CB1	TESTNOT03
90	2792452CB1	COLNTUT16
91	2853088CB1	PITUDIR01
92	2949004CB1	BRAITUT07
93	3011670CB1	PENITUT01
94	3242083CB1	FIBRTXS07
95	3363391CB1	PROSTUT10
96	3703614CB1	SINJNOT03
97	4000975CB1	HNT2AZS07
98	4598831CB1	UCMCNOT02

Table 5 (cont.)

Polynucleotide SEQ ID NO:	Incyte Project ID	Representative Library
99	4992201CB1	HNT2AGT01
100	5441583CB1	BSTMNON02
101	1639243CB1	UTRSNOT06
102	1335166CB1	COLNNOT13
103	166894CB1	LIVRNOT01
104	217969CB1	LATRTUT02
105	335237CB1	EOSIHET02
106	938306CB1	CERVNOT01
107	1448129CB1	PLACNOT02
108	1761049CB1	PITUNOT03
109	1959587CB1	HEAONOT03
110	2303463CB1	BRSTNOT05
111	2512281CB1	CONUTUT01
112	2755924CB1	THP1AZS08
113	2796369CB1	LIVRTUS02
135	3010920CB1..comp	MUSCNOT07
115	3360955CB1	PROSBPT02
136	3409459CB1..comp	STOMTUT02
117	4102938CB1	BRSTTUT17
118	4124601CB1	BRSTTUT26
119	4180577CB1	BRAIFEE04
120	5265807CB1	UTRENON03
121	5405979CB1	TESTNOT03
123	6247114CB1	TESTNOT17
124	3243866CB1	BRAIFEN03
125	7475633CB1	EPIMUNN04
126	1431268CB1	SINTBST01
127	2414185CB1	BRSTNOT04
128	5266594CB1	STOMFET01
129	7610617CB1	KIDCTME01
130	1902436CB1	OVARNOT07
131	2310369CB1	FIBRTXS07

Table 5 (cont.)

Polynucleotide SEQ ID NO:	Incyte Project ID	Representative Library
132	6180576CB1	HIPONOT01
133	2274523CB1	PROSNON01
134	1801820CB1	COLNNOT27

Table 6

Library	Vector	Library Description
ADRETUR01	PCDNA2.1	This random primed library was constructed using RNA isolated from left upper pole, adrenal gland tumor tissue removed from a 52-year-old Caucasian male during nephroureterectomy and local destruction of renal lesion. Pathology indicated grade 3 adrenal cortical carcinoma forming a mass that infiltrated almost the whole adrenal parenchyma and extended to adjacent adipose tissue. A metastatic tumor nodule was identified in the hilar region. The renal vein was infiltrated by tumor and the neoplastic process was present at the resection margin of the renal vein. Fragments of adrenal cortical carcinoma and thrombus were found in the inferior vena cava. Patient history included abnormal weight loss. Family history included skin cancer, type I diabetes, and neurotic depression.
BONSTUT01	pINCY	Library was constructed using RNA isolated from sacral bone tumor tissue removed from an 18-year-old Caucasian female during an exploratory laparotomy with soft tissue excision. Pathology indicated giant cell tumor of the sacrum. Patient history included a soft tissue malignant neoplasm. Family history included prostate cancer.
BRADDIR01	pINCY	Library was constructed using RNA isolated from diseased choroid plexus tissue of the lateral ventricle, removed from the brain of a 57-year-old Caucasian male, who died from a cerebrovascular accident.
BRAGNON02	pINCY	This normalized substantia nigra tissue library was constructed from 4.2x10e7 independent clones from a substantia nigra tissue library. Starting RNA was made from RNA isolated from substantia nigra tissue removed from an 81-year-old Caucasian female who died from a hemorrhage and ruptured thoracic aorta due to atherosclerosis. Pathology indicated moderate atherosclerosis involving the internal carotids, bilaterally; microscopic infarcts of the frontal cortex and hippocampus; and scattered diffuse amyloid plaques and neurofibrillary tangles, consistent with age. Grossly, the leptomeninges showed only mild thickening and hyalinization along the superior sagittal sinus. The remainder of the leptomeninges was thin and contained some congested blood vessels. Mild atrophy was found mostly in the frontal poles and lobes, and temporal lobes, bilaterally. Microscopically, there were pairs of Alzheimer type II astrocytes within the deep layers of the neocortex. There was increased satellitosis around neurons in the deep gray matter in the middle frontal cortex. The amygdala contained rare diffuse plaques and neurofibrillary tangles. The posterior hippocampus contained a

Table 6 (cont.)

Library	Vector	Library Description
		microscopic area of cystic cavitation with hemosiderin-laden macrophages surrounded by reactive gliosis. Patient history included sepsis, cholangitis, post-operative atelectasis, pneumonia CAD, cardiomegaly due to left ventricular hypertrophy, splenomegaly, arteriolonephrosclerosis, nodular colloidal goiter, emphysema, CHF, hypothyroidism, and peripheral vascular disease. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
BRAHTDR04	PCDNA2.1	This random primed library was constructed using RNA isolated archaecortex, anterior hippocampus tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated mild meningeal fibrosis predominately over the convexities, scattered axonal spheroids in the white matter of the cingulate cortex and the thalamus, and a few scattered neurofibrillary tangles in the entorhinal cortex and the periaqueductal gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with residual or relapsed tumor. Patient history included cholangiocarcinoma, post-operative Budd-Chiari syndrome, biliary ascites, hydrothorax, dehydration, malnutrition, oliguria and acute renal failure. Previous surgeries included cholecystectomy and resection of 85% of the liver.
BRAIFEE04	pINCY	This 5' biased random primed library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus who was stillborn with a hypoplastic left heart at 23 weeks' gestation.
BRAIFEN03	pINCY	This normalized fetal brain tissue library was constructed from 3.26 million independent clones from a fetal brain library. Starting RNA was made from brain tissue removed from a Caucasian male fetus, who was stillborn with a hypoplastic left heart at 23 weeks' gestation. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
BRAINON01	PSPORT1	This normalized library was made from 4.88 million independent clones from a brain tissue library. Starting RNA was isolated from brain tissue removed from a 26-year-old Caucasian male during cranioplasty and excision of a cerebral meningeal lesion. Pathology for the associated tumor tissue indicated a grade 4

Table 6 (cont.)

Library	Vector	Library Description
		oligoastrocytoma in the right fronto-parietal part of the brain.
BRAINOT09	pINCY	Library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus, who died at 23 weeks' gestation.
BRAITUT07	pINCY	Library was constructed using RNA isolated from left frontal lobe tumor tissue removed from the brain of a 32-year-old Caucasian male during excision of a cerebral meningeal lesion. Pathology indicated low grade desmoplastic neuronal neoplasm, type not otherwise specified. The lesion formed a firm, circumscribed cyst-associated mass involving white matter and cortex. No definite glial component was evident to suggest a diagnosis of ganglioglioma. Family history included atherosclerotic coronary artery disease.
BRSTNOT04	PSPORT1	Library was constructed using RNA isolated from breast tissue removed from a 62-year-old East Indian female during a unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated an invasive grade 3 ductal carcinoma. Patient history included benign hypertension, hyperlipidemia, and hematuria. Family history included cerebrovascular and cardiovascular disease, hyperlipidemia, and liver cancer.
BRSTNOT05	PSPORT1	Library was constructed using RNA isolated from breast tissue removed from a 58-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated multicentric invasive grade 4 lobular carcinoma. Patient history included skin cancer, rheumatic heart disease, osteoarthritis, and tuberculosis. Family history included cerebrovascular and cardiovascular disease, breast and prostate cancer, and type I diabetes.
BRSTTUT01	PSPORT1	Library was constructed using RNA isolated from breast tumor tissue removed from a 55-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated invasive grade 4 mammary adenocarcinoma of mixed lobular and ductal type, extensively involving the left breast. The tumor was identified in the deep dermis near the lactiferous ducts with extracapsular extension. Seven mid and low and five high axillary lymph nodes were positive for tumor. Proliferative fibrocysytic changes were characterized by apocrine metaplasia, sclerosing adenosis, cyst formation, and ductal hyperplasia without atypia. Patient history included atrial tachycardia, blood in the stool, and a benign breast neoplasm. Family history included benign hypertension, atherosclerotic coronary artery disease, cerebrovascular disease, and depressive disorder.

Table 6 (cont.)

Library	Vector	Library Description
BRSTTUT17	pINCY	Library was constructed using RNA isolated from left breast tumor tissue removed from a 65-year-old Caucasian female during a unilateral radical mastectomy. Pathology indicated invasive and in situ grade 3, nuclear grade 2 ductal carcinoma. Patient history included hyperlipidemia and uterine leiomyoma. Family history included stomach cancer, myocardial infarction, atherosclerotic coronary artery disease, prostate cancer, benign hypertension, breast cancer, and hyperlipidemia.
BRSTTUT26	pINCY	Library was constructed using RNA isolated from breast tumor tissue removed from an adult female. The breast carcinoma tumor tissue was found to have low vascular density and was considered resting.
BSTMNON02	PSPORT1	This normalized brain stem library was constructed from 2.84 million independent clones from a brain stem library. Starting RNA was made from the brain stem tissue of a 72-year-old Caucasian male who died from myocardial infarction. Patient history included coronary artery disease, insulin-dependent diabetes mellitus, and arthritis. Normalization and hybridization conditions were adapted from Soares et al. (PNAS (1994) 91:9228).
CERVNOT01	PSPORT1	Library was constructed using RNA isolated from the uterine cervical tissue of a 35-year-old Caucasian female during a vaginal hysterectomy with dilation and curettage. Pathology indicated mild chronic cervicitis. Family history included atherosclerotic coronary artery disease and type II diabetes.
COLNNOT13	pINCY	Library was constructed using RNA isolated from ascending colon tissue of a 28-year-old Caucasian male with moderate chronic ulcerative colitis.
COLNNOT27	pINCY	Library was constructed using RNA isolated from diseased cecal tissue removed from 31-year-old Caucasian male during a total intra-abdominal colectomy, appendectomy, and permanent ileostomy. Pathology indicated severe active Crohn's disease involving the colon from the cecum to the rectum. There were deep rake-like ulcerations which spared the intervening mucosa. The ulcers extended into the muscularis, and there was transmural inflammation. Patient history included an irritable colon. Previous surgeries included a colonscopy.
COLNTUT03	pINCY	Library was constructed using RNA isolated from colon tumor tissue obtained from the sigmoid colon of a 62-year-old Caucasian male during a sigmoidectomy and permanent colostomy. Pathology indicated invasive grade 2 adenocarcinoma. One lymph node contained metastasis with extranodal extension. Patient history

Table 6 (cont.)

Library	Vector	Library Description
		included hyperlipidemia, cataract disorder, and dermatitis. Family history included benign hypertension, atherosclerotic coronary artery disease, hyperlipidemia, breast cancer , and prostate cancer.
COLNTUT16	pINCY	Library was constructed using RNA isolated from colon tumor tissue obtained from a 60-year-old Caucasian male during a left hemicolectomy. Pathology indicated an invasive grade 2 adenocarcinoma, forming a sessile mass. Patient history included thrombophlebitis, inflammatory polyarthropathy, prostatic inflammatory disease, and depressive disorder. Previous surgeries included resection of the rectum. Family history included atherosclerotic coronary artery disease and colon cancer.
CONUTUT01	pINCY	Library was constructed using RNA isolated from sigmoid mesentery tumor tissue obtained from a 61-year-old female during a total abdominal hysterectomy and bilateral salpingo-oophorectomy with regional lymph node excision. Pathology indicated a metastatic grade 4 malignant mixed mullerian tumor present in the sigmoid mesentery at two sites.
CORPNOT02	pINCY	Library was constructed using RNA isolated from diseased corpus callosum tissue removed from the brain of a 74-year-old Caucasian male who died from Alzheimer's disease.
EOSIHET02	PBLUESCRIPT	Library was constructed using RNA isolated from peripheral blood cells apheresed from a 48-year-old Caucasian male. Patient history included hypereosinophilia. The cell population was determined to be greater than 77% eosinophils by Wright's staining.
EPIMUNN04	pINCY	This normalized mammary epithelial cell tissue library was constructed from 3.28 million independent clones from an epithelial cell tissue library. Starting RNA was made from untreated mammary epithelial cell tissue removed from a 21-year-old female. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
FIBRTXS07	pINCY	This subtracted library was constructed using 1.3 million clones from a dermal fibroblast library and was subjected to two rounds of subtraction hybridization with 2.8 million clones from the an untreated dermal fibroblast tissue library. The starting library for subtraction was constructed using RNA isolated from treated dermal fibroblast tissue removed from the breast of a 31-year-old

Table 6 (cont.)

Library	Vector	Library Description
		Caucasian female. The cells were treated with 9CIS retinoic acid. The hybridization probe for subtraction was derived from a similarly constructed library from RNA isolated from untreated dermal fibroblast tissue from the same donor. Subtractive hybridization conditions were based on the methodologies of Swaroop et al., NAR (1991) 19:1954 and Bonaldo, et al., Genome Research (1996) 6:791.
HEAONOT03	pINCY	Library was constructed using RNA isolated from aortic tissue removed from a 27-year-old Caucasian female, who died from an intracranial bleed.
HIPONOT01	PBLUESCRIPT	Library was constructed using RNA isolated from the hippocampus tissue of a 72-year-old Caucasian female who died from an intracranial bleed. Patient history included nose cancer, hypertension, and arthritis.
HNT2AGT01	PBLUESCRIPT	Library was constructed at Stratagene (STR937233), using RNA isolated from the hNT2 cell line derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor. Cells were treated with retinoic acid for 5 weeks and with mitotic inhibitors for two weeks and allowed to mature for an additional 4 weeks in conditioned medium.
HNT2AZS07	PSPORT1	This subtracted library was constructed from RNA isolated from an hNT2 cell line (derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor) treated for three days with 0.35 micromolar AZ. The hybridization probe for subtraction was derived from a similarly constructed library from untreated hNT2 cells. 3.08M clones from the AZ-treated library were subjected to three rounds of subtractive hybridization with 3.04M clones from the untreated library. Subtractive hybridization conditions were based on the methodologies of Swaroop et al. (NAR (1991) 19:1954) and Bonaldo et al. (Genome Research (1996) 6:791).
KERANOT01	PBLUESCRIPT	Library was constructed using RNA isolated from neonatal keratinocytes obtained from the leg skin of a spontaneously aborted black male.
KIDCTME01	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from kidney cortex tissue removed from a 65-year-old male during nephroureterectomy. Pathology indicated the margins of resection were free of involvement. Pathology for the matched tumor tissue indicated grade 3 renal cell carcinoma, clear cell type, forming a variegated multicystic mass situated within the mid-portion of the kidney. The tumor invaded deeply into but not through the renal capsule.

Table 6 (cont.)

Library	Vector	Library Description
KIDNNOT09	pINCY	Library was constructed using RNA isolated from the kidney tissue of a Caucasian male fetus, who died at 23 weeks' gestation.
KIDNNOT19	pINCY	Library was constructed using RNA isolated from kidney tissue removed a 65-year-old Caucasian male during an exploratory laparotomy and nephroureterectomy. Pathology for the associated tumor tissue indicated a grade 1 renal cell carcinoma within the upper pole of the left kidney. Patient history included malignant melanoma of the abdominal skin, benign neoplasm of colon, cerebrovascular disease, and umbilical hernia. Family history included myocardial infarction, atherosclerotic coronary artery disease, cerebrovascular disease, prostate cancer, myocardial infarction, and atherosclerotic coronary artery disease.
LATRTUT02	pINCY	Library was constructed using RNA isolated from a myxoma removed from the left atrium of a 43-year-old Caucasian male during annuloplasty. Pathology indicated atrial myxoma. Patient history included pulmonary insufficiency, acute myocardial infarction, atherosclerotic coronary artery disease, hyperlipidemia, and tobacco use. Family history included benign hypertension, acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
LIVRNON08	pINCY	This normalized library was constructed from 5.7 million independent clones from a pooled liver tissue library. Starting RNA was made from pooled liver tissue removed from a 4-year-old Hispanic male who died from anoxia and a 16 week female fetus who died after 16-weeks gestation from anencephaly. Serologies were positive for cytolomegalovirus in the 4-year-old. Patient history included asthma in the 4-year-old. Family history included taking daily prenatal vitamins and mitral valve prolapse in the mother of the fetus. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
LIVRNOT01	PBLUESCRIPT	Library was constructed at Stratagene, using RNA isolated from the liver tissue of a 49-year-old male.
LIVRTUS02	pINCY	This subtracted C3A liver tumor cell line tissue library was constructed using 6.4 million clones from a 3-metnylcholthrene-treated hepatocyte library and was subjected to two rounds of subtraction hybridization with 1.72 million clones from an untreated C3A hepatocyte library. The starting library for subtraction was constructed using RNA isolated from a treated C3A hepatocyte cell line which is a

Table 6 (cont.)

Library	Vector	Library Description
		derivative of Hep G2, a cell line derived from a hepatoblastoma removed from a 15-year-old Caucasian male. The cells were treated with 3-methylcholanthrene (MCA), 5 mM for 48 hours. The hybridization probe for subtraction was derived from a similarly constructed library from RNA isolated from untreated C3A hepatocyte cells from the same cell line. Subtractive hybridization conditions were based on the methodologies of Swaroop et al., NAR 19 (1991):1954 and Bonaldo, et al. Genome Research 6(1996):791.
LVENNOT01	PBLUESCRIPT	Library was constructed using RNA isolated from the left ventricle of a 51-year-old Caucasian female, who died from an intracranial bleed.
MONOTXS05	pINCY	Library was constructed using 7.5 million clones from a treated monocyte library and subjected to two rounds of subtraction hybridization with 1.03 x 10e7 clones from a second treated monocyte library. The starting library for subtraction was constructed using RNA from treated monocytes from peripheral blood obtained from a 42-year-old female. The cells were pre-treated for 1 hour with 10 ng/ml anti-interleukin-10 (anti-IL-10). Lipopolysaccharide (LPS) at 5 ng/ml was added and monocytes harvested after 24 hours. Monocytes were isolated from buffy coat by adherence to plastic. The hybridization probe for subtraction was derived from a similarly constructed library using RNA isolated from monocyte tissue, treated with interleukin-10 (IL10) and lipopolysaccharide (LPS) from the same donor. Subtractive hybridization conditions were based on the methodologies of Swaroop et al. NAR 19 (1991):1954 and Bonaldo, et al. Genome Research 6 (1996):791.
MUSCNOT07	pINCY	Library was constructed using RNA isolated from muscle tissue removed from the forearm of a 38-year-old Caucasian female during a soft tissue excision. Pathology for the associated tumor tissue indicated intramuscular hemangioma. Family history included breast cancer, benign hypertension, cerebrovascular disease, colon cancer, and type II diabetes.
OVARNOT07	pINCY	Library was constructed using RNA isolated from left ovarian tissue removed from a 28-year-old Caucasian female during a vaginal hysterectomy and removal of the fallopian tubes and ovaries. The tissue was associated with multiple follicular cysts, endometrium in a weakly proliferative phase, and chronic cervicitis of the cervix with squamous metaplasia. Family history included benign hypertension, hyperlipidemia, and atherosclerotic coronary artery disease.

Table 6 (cont.)

Library	Vector	Library Description
OVARTUE01	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from left ovary tumor tissue removed from a 44-year-old female. Pathology indicated grade 4 (of 4) serous carcinoma replacing both the right and left ovaries forming solid mass cystic masses. Neoplastic deposits are identified in para-ovarian soft tissue, on the surface of the uterus and scattered throughout the myometrium and cervix. Transverse colon was positive for metastatic disease. Multiple staging biopsies including diaphragm, bladder, liver, sigmoid rectal serosa, inguinal lymph nodes and left colic gutter are metastatically involved.
PENITUT01	pINCY	Library was constructed using RNA isolated from tumor tissue removed from the penis of a 64-year-old Caucasian male during penile amputation. Pathology indicated a fungating invasive grade 4 squamous cell carcinoma involving the inner wall of the foreskin and extending onto the glans penis. Patient history included benign neoplasm of the large bowel, atherosclerotic coronary artery disease, angina pectoris, gout, and obesity. Family history included malignant pharyngeal neoplasm, chronic lymphocytic leukemia, and chronic liver disease.
PGANNON02	PSPORT1	This normalized paraganglion library was constructed with 5.48e6 independent clones from a paraganglionic tissue library. Starting RNA was made from paraganglionic tissue removed from a 46-year-old Caucasian male during exploratory laparotomy. Pathology indicated a benign paraganglioma and was associated with a grade 2 renal cell carcinoma. The normalization and hybridization conditions were adapted from Soares et al. (PNAS (1994) 91:9228-9232) using a significantly longer (48-hour) reannealing hybridization period.
PITUDIR01	PCDNA2.1	This random primed library was constructed using RNA isolated from pituitary gland tissue removed from a 70-year-old female who died from metastatic adenocarcinoma. Pathology for the brain indicated moderate Alzheimer disease and mild carotid and cerebral atherosclerosis. The cerebral hemispheres, frontal and temporal lobes, white matter, and hippocampus showed mild atrophy, bilaterally. There were numerous neurofibrillary tangles, neuritic and diffuse amyloid plaques deposited throughout most neocortical areas. Most of the diffuse plaques were in the superficial layers, with more core and neuritic amyloid plaques in the deep cortical layers. Most of the tangles were found in small interneurons, rather than in the large pyramidal neurons. The areas that were most involved with plaques and tangles were the entorhinal cortex, temporal cortex, and superior parietal lobes.

Table 6 (cont.)

Library	Vector	Library Description
		There was marked vacuolization of the superficial layers throughout all neocortical areas examined. The hippocampus contained numerous neurofibrillary tangles (predominantly in the CA-1 field), diffuse and neuritic plaques, as well as granulovacuolar degeneration within the pyramidal cell neurons. There were neuritic plaques with scattered neurofibrillary tangles within the amygdala. The thalamus had scattered diffuse plaques. There was mild pigment incontinence in the substantia nigra compacta. The periaqueductal gray matter showed mild gliosis. Diffuse plaques were found within the superior colliculus. Neurofibrillary tangles were found within the pons. The neurons of the locus ceruleus were ballooned and contain eosinophilic foamy material with very little neuromelanin pigment.
PITUNOT03	PSPORT1	Library was constructed using RNA isolated from pituitary tissue of a 46-year-old Caucasian male, who died from colon cancer. Serologies were negative. Patient history included arthritis, peptic ulcer disease, and tobacco use. Patient medications included Tagamet and muscle relaxants.
PLACNOB01	PBLUESCRIPT	Library was constructed using RNA isolated from placenta.
PLACNOT02	pINCY	Library was constructed using RNA isolated from the placental tissue of a Hispanic female fetus, who was prematurely delivered at 21 weeks' gestation. Serologies of the mother's blood were positive for CMV (cytomegalovirus).
PROSBPT02	pINCY	Library was constructed using RNA isolated from diseased prostate tissue removed from a 65-year-old Caucasian male during a radical prostatectomy. Pathology indicated benign prostatic hyperplasia (BPH). Pathology for the associated tumor tissue indicated adenocarcinoma, Gleason grade 3+4. One (of 7) right pelvic lymph nodes was positive for metastatic adenocarcinoma. The patient presented with induration and elevated prostate specific antigen (PSA). Patient history included a benign neoplasm of the large bowel and benign hypertension.
PROSNON01	PSPORT1	This normalized prostate library was constructed from 4.4 M independent clones from a prostate library. Starting RNA was made from prostate tissue removed from a 28-year-old Caucasian male who died from a self-inflicted gunshot wound. The normalization and hybridization conditions were adapted from Soares, M.B. et al. (1994) Proc. Natl. Acad. Sci. USA 91:9228-9232, using a longer (19 hour) reannealing hybridization period.

Table 6 (cont.)

Library	Vector	Library Description
PROSNOT19	pINCY	Library was constructed using RNA isolated from diseased prostate tissue removed from a 59-year-old Caucasian male during a radical prostatectomy with regional lymph node excision. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated an adenocarcinoma (Gleason grade 3+3). The patient presented with elevated prostate-specific antigen (PSA). Patient history included colon diverticuli, asbestosis, and thrombophlebitis. Previous surgeries included a partial colectomy. Family history included benign hypertension, multiple myeloma, hyperlipidemia and rheumatoid arthritis.
PROSTUT10	pINCY	Library was constructed using RNA isolated from prostatic tumor tissue removed from a 66-year-old Caucasian male during radical prostatectomy and regional lymph node excision. Pathology indicated an adenocarcinoma (Gleason grade 2+3). Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA). Family history included prostate cancer and secondary bone cancer.
SINJNOT03	pINCY	Library was constructed using RNA isolated from duodenum tissue removed from the small intestine of a 16-year-old Caucasian male who died from head trauma. Patient history included a kidney infection.
SINTBST01	pINCY	Library was constructed using RNA isolated from ileum tissue obtained from an 18-year-old Caucasian female during bowel anastomosis. Pathology indicated Crohn's disease of the ileum, involving 15 cm of the small bowel. Family history included cerebrovascular disease and atherosclerotic coronary artery disease.
STOMFET01	pINCY	Library was constructed using RNA isolated from the stomach tissue of a Caucasian female fetus, who died at 20 weeks' gestation.
STOMTUT02	pINCY	Library was constructed using RNA isolated from stomach tumor tissue obtained from a 68-year-old Caucasian female during a partial gastrectomy. Pathology indicated a malignant lymphoma of diffuse large-cell type. Previous surgeries included cholecystectomy. Patient history included thalassemia. Family history included acute leukemia, malignant neoplasm of the esophagus, malignant stomach neoplasm, and atherosclerotic coronary artery disease.
TESTNOC01	PBLUESCRIPT	This large size fractionated library was constructed using RNA isolated from testicular tissue removed from a pool of eleven, 10 to 61-year-old Caucasian males.

Table 6 (cont.)

Library	Vector	Library Description
TESTNOT03	PBLUESCRIPT	Library was constructed using RNA isolated from testicular tissue removed from a 37-year-old Caucasian male, who died from liver disease. Patient history included cirrhosis, jaundice, and liver failure.
TESTNOT17	pINCY	Library was constructed from testis tissue removed from a 26-year-old Caucasian male who died from head trauma due to a motor vehicle accident. Serologies were negative. Patient history included a hernia at birth, tobacco use (1 1/2 ppd), marijuana use, and daily alcohol use (beer and hard liquor).
THP1AZS08	PSPORT1	This subtracted THP-1 promonocyte cell line library was constructed using 5.76 x 1e6 clones from a 5-aza-2'-deoxycytidine (AZ) treated THP-1 cell library. Starting RNA was made from THP-1 promonocyte cells treated for three days with 0.8 micromolar AZ. The donor had acute monocytic leukemia The hybridization probe for subtraction was derived from a similarly constructed library, made from 1 microgram of polyA RNA isolated from untreated THP-1 cells. 5.76 million clones from the AZ-treated THP-1 cell library were then subjected to two rounds of subtractive hybridization with 5 million clones from the untreated THP-1 cell library. Subtractive hybridization conditions were based on the methodologies of Swaroop et al., NAR (1991) 19:1954, and Bonaldo et al., Genome Research (1996) 6:791.
UCMCNOT02	pINCY	Library was constructed using RNA isolated from mononuclear cells obtained from the umbilical cord blood of nine individuals.

Table 6 (cont.)

Library	Vector	Library Description
UTRENON03	pINCY	Library was constructed from 1.2 x 10e7 independent clones from a uterine endometrial tissue library. Starting RNA was made from uterine endometrium tissue obtained from a 29-year-old Caucasian female during a vaginal hysterectomy and cystocele repair. Pathology indicated the endometrium was secretory and the cervix showed mild chronic cervicitis with focal squamous metaplasia. Pathology for the associated tumor tissue indicated an intramural uterine leiomyoma. Patient history included hypothyroidism, pelvic floor relaxation, incomplete T-12 injury (due to a motor vehicle accident) causing paraplegia and self catheterization. Previous surgeries included a normal delivery, a laminectomy, and a rhinoplasty. Family history included benign hypertension, type II diabetes, and hyperlipidemia. The libraries were normalized in two rounds using conditions adapted from Soares et a 1., PNAS (1994) 91:9928 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
UTRSNOT06	pINCY	Library was constructed using RNA isolated from myometrial tissue removed from a 50-year-old Caucasian female during a vaginal hysterectomy. Pathology indicated residual atypical complex endometrial hyperplasia. Pathology for the associated tissue removed during dilation and curettage indicated fragments of atypical complex hyperplasia and a single microscopic focus suspicious for grade 1 adenocarcinoma. Patient history included benign breast neoplasm, hypothyroid disease, polypectomy, and arthralgia. Family history included cerebrovascular disease, atherosclerotic coronary artery disease, hyperlipidemia, and chronic hepatitis.

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences*: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs*: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	*PFAM hits:* Probability value= 1.0E-3 or less *Signal peptide hits:* Score= 0 or greater

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score≥GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:

a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-67,

b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-67,

c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, and

d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67.

2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-67.

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:68-134.

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

9. A method of producing a polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

10. A method of claim 9, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-67.

11. An isolated antibody which specifically binds to a polypeptide of claim 1.

12. An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:68-134,

b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:68-134,

c) a polynucleotide complementary to a polynucleotide of a),

d) a polynucleotide complementary to a polynucleotide of b), and

e) an RNA equivalent of a)-d).

13. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 12.

14. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

15. A method of claim 14, wherein the probe comprises at least 60 contiguous nucleotides.

16. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and

b) detecting the presence or absence of said amplified target polynucleotide or fragment

thereof, and, optionally, if present, the amount thereof.

17. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

18. A composition of claim 17, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-67.

19. A method for treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment the composition of claim 17.

20. A method of screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

21. A composition comprising an agonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

22. A method for treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment a composition of claim 21.

23. A method of screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting antagonist activity in the sample.

24. A composition comprising an antagonist compound identified by a method of claim 23 and a pharmaceutically acceptable excipient.

25. A method for treating a disease or condition associated with overexpression of functional SECP, comprising administering to a patient in need of such treatment a composition of claim 24.

26. A method of screening for a compound that specifically binds to the polypeptide of claim 1, the method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and

b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

27. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, the method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,

b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and

c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

28. A method of screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,

b) detecting altered expression of the target polynucleotide, and

c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

29. A method of assessing toxicity of a test compound, the method comprising:

a) treating a biological sample containing nucleic acids with the test compound,

b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 12 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 12 or fragment thereof,

c) quantifying the amount of hybridization complex, and

d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

30. A diagnostic test for a condition or disease associated with the expression of SECP in a biological sample, the method comprising:

a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex, and

b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

31. The antibody of claim 11, wherein the antibody is:

a) a chimeric antibody,

b) a single chain antibody,

c) a Fab fragment,

d) a $F(ab')_2$ fragment, or

e) a humanized antibody.

32. A composition comprising an antibody of claim 11 and an acceptable excipient.

33. A method of diagnosing a condition or disease associated with the expression of SECP in a subject, comprising administering to said subject an effective amount of the composition of claim 32.

34. A composition of claim 32, wherein the antibody is labeled.

35. A method of diagnosing a condition or disease associated with the expression of SECP in a subject, comprising administering to said subject an effective amount of the composition of claim 34.

36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11, the method comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, or an immunogenic fragment thereof, under conditions to elicit an antibody response,

b) isolating antibodies from said animal, and

c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67.

37. A polyclonal antibody produced by a method of claim 36.

38. A composition comprising the polyclonal antibody of claim 37 and a suitable carrier.

39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11, the method comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67, or an immunogenic fragment thereof, under conditions to elicit an antibody response,

b) isolating antibody producing cells from the animal,

c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,

d) culturing the hybridoma cells, and

e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67.

40. A monoclonal antibody produced by a method of claim 39.

41. A composition comprising the monoclonal antibody of claim 40 and a suitable carrier.

42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.

43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant immunoglobulin library.

44. A method of detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67 in a sample, the method comprising:

a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and

b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67 in the sample.

45. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67 from a sample, the method comprising:

a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and

b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-67.

46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 13.

47. A method of generating a transcript image of a sample which contains polynucleotides, the method comprising:

a) labeling the polynucleotides of the sample,

b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and

c) quantifying the expression of the polynucleotides in the sample.

48. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.

49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to said target polynucleotide.

52. An array of claim 48, which is a microarray.

53. An array of claim 48, further comprising said target polynucleotide hybridized to a nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

55. An array of claim 48, wherein each distinct physical location on the substrate contains multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical location have the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another distinct physical location on the substrate.

56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.

57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.

66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.

67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.

68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.

69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.

70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.

71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.

72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.

73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.

74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.

75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.

76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:21.

77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:22.

78. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:23.

79. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:24.

80. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:25.

81. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:26.

82. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:27.

83. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:28.

84. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:29.

85. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:30.

86. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:31.

87. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:32.

88. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:33.

89. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:34.

90. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:35.

91. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:36.

92. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:37.

93. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:38.

94. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:39.

95. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:40.

96. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:41.

97. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:42.

98. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:43.

99. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:44.

100. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:45.

101. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:46.

102. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:47.

103. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:48.

104. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:49.

105. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:50.

106. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:51.

107. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:52.

108. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:53.

109. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:54.

110. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:55.

111. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:56.

112. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:57.

113. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:58.

114. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:59.

115. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:60.

116. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:61.

117. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:62.

118. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:63.

119. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:64.

120. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:65.

121. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:66.

122. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:67.

123. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:68.

124. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:69.

125. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:70.

126. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:71.

127. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:72.

128. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:73.

129. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:74.

130. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:75.

131. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID

NO:76.

132. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:77.

133. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:78.

134. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:79.

135. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:80.

136. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:81.

137. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:82.

138. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:83.

139. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:84.

140. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:85.

141. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:86.

142. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:87.

143. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:88.

144. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:89.

145. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:90.

146. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:91.

147. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:92.

148. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:93.

149. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:94.

150. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:95.

151. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:96.

152. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:97.

153. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:98.

154. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:99.

155. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:100.

156. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:101.

157. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:102.

158. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:103.

159. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:104.

160. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:105.

161. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:106.

162. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:107.

163. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:108.

164. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:109.

165. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:110.

166. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:111.

167. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:112.

168. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:113.

169. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:114.

170. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:115.

171. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:116.

172. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:117.

173. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:118.

174. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:119.

175. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:120.

176. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:121.

177. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:122.

178. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:123.

179. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:124.

180. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:125.

181. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:126.

182. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:127.

183. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:128.

184. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:129.

185. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:130.

186. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:131.

187. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:132.

188. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:133.

189. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:134.

<110> INCYTE GENOMICS, INC.
YUE, Henry
TANG, Y. Tom
NGUYEN, Danniel B.
YAO, Monique G.
XU, Yuming
TRIBOULEY, Catherine M.
SANJANWALA, Madhu S.
WALIA, Narinder K.
BAUGHN, Mariah R.
SAPPERSTEIN, Stephanie
LAL, Preeti
THORNTON, Michael
GANDHI, Ameena R.
RAMKUMAR, Jayalaxmi
ELLIOTT, Vicki S.
ARVIZU, Chandra
THANGAVELU, Kavitha
GIETZEN, Kimberly
DING, Li
AU-YOUNG, Janice
TRAN, Bao
POLICKY, Jennifer L.
LEE, Sally
LU, Dyung Aina M.
BURFORD, Neil
WARREN, Bridget A.
GURURAJAN, Rajagopal
DUGGAN, Brendan M.
HONCHELL, Cynthia D.
HAFALIA, April J.A.

<120> SECRETED PROTEINS

<130> PI-0240 PCT

<140> To Be Assigned
<141> Herewith

<150> 60/236,869; 60/239,812; 60/240,108; 60/241,282; 60/242,218
<151> 2000-09-29; 2000-10-11; 2000-10-12; 2000-10-17; 2000-10-20

<160> 136

<170> PERL Program

<210> 1
<211> 622
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3211795CD1

<400> 1
Met Ala Ala Ala Pro Arg Ala Gly Arg Arg Arg Gly Gln Pro Leu
1 5 10 15
Leu Ala Leu Leu Leu Leu Leu Leu Ala Pro Leu Pro Pro Gly Ala
20 25 30
Pro Pro Gly Ala Asp Ala Tyr Phe Pro Glu Glu Arg Trp Ser Pro
35 40 45
Glu Ser Pro Leu Gln Ala Pro Arg Val Leu Ile Ala Leu Leu Ala

```
                    50                  55                  60
Arg Asn Ala Ala His Ala Leu Pro Thr Thr Leu Gly Ala Leu Glu
                    65                  70                  75
Arg Leu Arg His Pro Arg Glu Arg Thr Ala Leu Trp Val Ala Thr
                    80                  85                  90
Asp His Asn Met Asp Asn Thr Ser Thr Val Leu Arg Glu Trp Leu
                    95                  100                 105
Val Ala Val Lys Ser Leu Tyr His Ser Val Glu Trp Arg Pro Ala
                    110                 115                 120
Glu Glu Pro Arg Ser Tyr Pro Asp Glu Glu Gly Pro Lys His Trp
                    125                 130                 135
Ser Asp Ser Arg Tyr Glu His Val Met Lys Leu Arg Gln Ala Ala
                    140                 145                 150
Leu Lys Ser Ala Arg Asp Met Trp Ala Asp Tyr Ile Leu Phe Val
                    155                 160                 165
Asp Ala Asp Asn Leu Ile Leu Asn Pro Asp Thr Leu Ser Leu Leu
                    170                 175                 180
Ile Ala Glu Asn Lys Thr Val Val Ala Pro Met Leu Asp Ser Arg
                    185                 190                 195
Ala Ala Tyr Ser Asn Phe Trp Cys Gly Met Thr Ser Gln Gly Tyr
                    200                 205                 210
Tyr Lys Arg Thr Pro Ala Tyr Ile Pro Ile Arg Lys Arg Asp Arg
                    215                 220                 225
Arg Gly Cys Phe Ala Val Pro Met Val His Ser Thr Phe Leu Ile
                    230                 235                 240
Asp Leu Arg Lys Ala Ala Ser Arg Asn Leu Ala Phe Tyr Pro Pro
                    245                 250                 255
His Pro Asp Tyr Thr Trp Ser Phe Asp Asp Ile Ile Val Phe Ala
                    260                 265                 270
Phe Ser Cys Lys Gln Ala Glu Val Gln Met Tyr Val Cys Asn Lys
                    275                 280                 285
Glu Glu Tyr Gly Phe Leu Pro Val Pro Leu Arg Ala His Ser Thr
                    290                 295                 300
Leu Gln Asp Glu Ala Glu Ser Phe Met His Val Gln Leu Glu Val
                    305                 310                 315
Met Val Lys His Pro Pro Ala Glu Pro Ser Arg Phe Ile Ser Ala
                    320                 325                 330
Pro Thr Lys Thr Pro Asp Lys Met Gly Phe Asp Glu Val Phe Met
                    335                 340                 345
Ile Asn Leu Arg Arg Arg Gln Asp Arg Arg Glu Arg Met Leu Arg
                    350                 355                 360
Ala Leu Gln Ala Gln Glu Ile Glu Cys Arg Leu Val Glu Ala Val
                    365                 370                 375
Asp Gly Lys Ala Met Asn Thr Ser Gln Val Glu Ala Leu Gly Ile
                    380                 385                 390
Gln Met Leu Pro Gly Tyr Arg Asp Pro Tyr His Gly Arg Pro Leu
                    395                 400                 405
Thr Lys Gly Glu Leu Gly Cys Phe Leu Ser His Tyr Asn Ile Trp
                    410                 415                 420
Lys Glu Val Val Asp Arg Gly Leu Gln Lys Ser Leu Val Phe Glu
                    425                 430                 435
Asp Asp Leu Arg Phe Glu Ile Phe Phe Lys Arg Arg Leu Met Asn
                    440                 445                 450
Leu Met Arg Asp Val Glu Arg Glu Gly Leu Asp Trp Asp Leu Ile
                    455                 460                 465
Tyr Val Gly Arg Lys Arg Met Gln Val Glu His Pro Glu Lys Ala
                    470                 475                 480
Val Pro Arg Val Arg Asn Leu Val Glu Ala Asp Tyr Ser Tyr Trp
                    485                 490                 495
Thr Leu Ala Tyr Val Ile Ser Leu Gln Gly Ala Arg Lys Leu Leu
                    500                 505                 510
Ala Ala Glu Pro Leu Ser Lys Met Leu Pro Val Asp Glu Phe Leu
                    515                 520                 525
```

```
Pro Val Met Phe Asp Lys His Pro Val Ser Glu Tyr Lys Ala His
                530                 535                 540
Phe Ser Leu Arg Asn Leu His Ala Phe Ser Val Glu Pro Leu Leu
                545                 550                 555
Ile Tyr Pro Thr His Tyr Thr Gly Asp Asp Gly Tyr Val Ser Asp
                560                 565                 570
Thr Glu Thr Ser Val Val Trp Asn Asn Glu His Val Lys Thr Asp
                575                 580                 585
Trp Asp Arg Ala Lys Ser Gln Lys Met Arg Glu Gln Gln Ala Leu
                590                 595                 600
Ser Arg Glu Ala Lys Asn Ser Asp Val Leu Gln Ser Pro Leu Asp
                605                 610                 615
Ser Ala Ala Arg Asp Glu Leu
                620
```

<210> 2
<211> 529
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6813464CD1

<400> 2

```
Met Ala Gly Ala Pro Pro Pro Ala Leu Leu Leu Pro Cys Ser Leu
  1                 5                  10                  15
Ile Ser Asp Cys Cys Ala Ser Asn Gln Arg His Ser Val Gly Val
                 20                  25                  30
Gly Pro Ser Glu Leu Val Lys Lys Gln Ile Glu Leu Lys Ser Arg
                 35                  40                  45
Gly Val Lys Leu Met Pro Ser Lys Asp Asn Ser Gln Lys Thr Ser
                 50                  55                  60
Val Leu Thr Gln Val Gly Val Ser Gln Gly His Asn Met Cys Pro
                 65                  70                  75
Asp Pro Gly Ile Pro Glu Arg Gly Lys Arg Leu Gly Ser Asp Phe
                 80                  85                  90
Arg Leu Gly Ser Ser Val Gln Phe Thr Cys Asn Glu Gly Tyr Asp
                 95                 100                 105
Leu Gln Gly Ser Lys Arg Ile Thr Cys Met Lys Val Ser Asp Met
                110                 115                 120
Phe Ala Ala Trp Ser Asp His Arg Pro Val Cys Arg Ala Arg Met
                125                 130                 135
Cys Asp Ala His Leu Arg Gly Pro Ser Gly Ile Ile Thr Ser Pro
                140                 145                 150
Asn Phe Pro Ile Gln Tyr Asp Asn Asn Ala His Cys Val Trp Ile
                155                 160                 165
Ile Thr Ala Leu Asn Pro Ser Lys Val Ile Lys Leu Ala Phe Glu
                170                 175                 180
Glu Phe Asp Leu Glu Arg Gly Tyr Asp Thr Leu Thr Val Gly Asp
                185                 190                 195
Gly Gly Gln Asp Gly Asp Gln Lys Thr Val Leu Tyr Ile Leu Thr
                200                 205                 210
Gly Thr Ser Val Pro Asp Leu Ile Val Ser Thr Asn His Gln Met
                215                 220                 225
Trp Leu Leu Phe Gln Thr Asp Gly Ser Gly Ser Ser Leu Gly Phe
                230                 235                 240
Lys Ala Ser Tyr Glu Glu Ile Glu Gln Gly Ser Cys Gly Asp Pro
                245                 250                 255
Gly Ile Pro Ala Tyr Gly Arg Arg Glu Gly Ser Arg Phe His His
                260                 265                 270
Gly Asp Thr Leu Lys Phe Glu Cys Gln Pro Ala Phe Glu Leu Val
                275                 280                 285
```

```
Gly Gln Lys Ala Ile Thr Cys Gln Lys Asn Asn Gln Trp Ser Ala
                290                 295                 300
Lys Lys Pro Gly Cys Val Phe Ser Cys Phe Phe Asn Phe Thr Ser
                305                 310                 315
Pro Ser Gly Val Val Leu Ser Pro Asn Tyr Pro Glu Asp Tyr Gly
                320                 325                 330
Asn His Leu His Cys Val Trp Leu Ile Leu Ala Arg Pro Glu Ser
                335                 340                 345
Arg Ile His Leu Ala Phe Asn Asp Ile Asp Val Glu Pro Gln Phe
                350                 355                 360
Asp Phe Leu Val Ile Lys Asp Gly Ala Thr Ala Glu Ala Pro Val
                365                 370                 375
Leu Gly Thr Phe Ser Gly Asn Gln Leu Pro Ser Ser Ile Thr Ser
                380                 385                 390
Ser Gly His Val Ala Arg Leu Glu Phe Gln Thr Asp His Ser Thr
                395                 400                 405
Gly Lys Arg Gly Phe Asn Ile Thr Phe Thr Thr Phe Arg His Asn
                410                 415                 420
Glu Cys Pro Asp Pro Gly Val Pro Val Asn Gly Lys Arg Phe Gly
                425                 430                 435
Asp Ser Leu Gln Leu Gly Ser Ser Ile Ser Phe Leu Cys Asp Glu
                440                 445                 450
Gly Phe Leu Gly Thr Gln Gly Ser Glu Thr Ile Thr Cys Val Leu
                455                 460                 465
Lys Glu Gly Ser Val Val Trp Asn Ser Ala Val Leu Arg Cys Glu
                470                 475                 480
Ala Pro Cys Gly Gly His Leu Thr Ser Pro Ser Gly Thr Ile Leu
                485                 490                 495
Ser Pro Gly Trp Pro Gly Phe Tyr Lys Asp Ala Leu Ser Cys Ala
                500                 505                 510
Trp Val Ile Glu Ala Gln Pro Gly Tyr Pro Ile Lys Ile Thr Phe
                515                 520                 525
Asp Arg Cys Leu
```

<210> 3
<211> 204
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2156540CD1

<400> 3

```
Met Gly Ala Pro Leu Ala Val Ala Leu Gly Ala Leu His Tyr Leu
  1                 5                  10                  15
Ala Leu Phe Leu Gln Leu Gly Gly Ala Thr Arg Pro Ala Gly His
                 20                  25                  30
Ala Pro Trp Asp Asn His Val Ser Gly His Ala Leu Phe Thr Glu
                 35                  40                  45
Thr Pro His Asp Met Thr Ala Arg Thr Gly Glu Asp Val Glu Met
                 50                  55                  60
Ala Cys Ser Phe Arg Gly Ser Gly Ser Pro Ser Tyr Ser Leu Glu
                 65                  70                  75
Ile Gln Trp Trp Tyr Val Arg Ser His Arg Asp Trp Thr Asp Lys
                 80                  85                  90
Gln Ala Trp Ala Ser Asn Gln Leu Lys Ala Ser Gln Gln Glu Asp
                 95                 100                 105
Ala Gly Lys Glu Ala Thr Lys Ile Ser Val Val Lys Val Val Gly
                110                 115                 120
Ser Asn Ile Ser His Lys Leu Arg Leu Ser Arg Val Lys Pro Thr
                125                 130                 135
```

```
Asp Glu Gly Thr Tyr Glu Cys Arg Val Ile Asp Phe Ser Asp Gly
                  140                 145                 150
Lys Ala Arg His His Lys Val Lys Ala Tyr Leu Arg Val Gln Pro
                  155                 160                 165
Gly Glu Asn Ser Val Leu His Leu Pro Glu Ala Pro Pro Ala Ala
                  170                 175                 180
Pro Ala Pro Pro Pro Pro Lys Pro Gly Lys Glu Leu Arg Lys Arg
                  185                 190                 195
Ser Val Asp Gln Glu Ala Cys Ser Leu
                  200
```

<210> 4
<211> 406
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 894939CD1

```
<400> 4
Met Arg Gly Ser Val Glu Cys Thr Trp Gly Trp Gly His Cys Ala
  1                 5                  10                  15
Pro Ser Pro Leu Leu Leu Trp Thr Leu Leu Leu Phe Ala Ala Pro
                  20                  25                  30
Phe Gly Leu Leu Gly Glu Lys Thr Arg Gln Val Ser Leu Glu Val
                  35                  40                  45
Ile Pro Asn Trp Leu Gly Pro Leu Gln Asn Leu Leu His Ile Arg
                  50                  55                  60
Ala Val Gly Thr Asn Ser Thr Leu His Tyr Val Trp Ser Ser Leu
                  65                  70                  75
Gly Pro Leu Ala Val Val Met Val Ala Thr Asn Thr Pro His Ser
                  80                  85                  90
Thr Leu Ser Val Asn Trp Ser Leu Leu Leu Ser Pro Glu Pro Asp
                  95                 100                 105
Gly Gly Leu Met Val Leu Pro Lys Asp Ser Ile Gln Phe Ser Ser
                 110                 115                 120
Ala Leu Val Phe Thr Arg Leu Leu Glu Phe Asp Ser Thr Asn Val
                 125                 130                 135
Ser Asp Thr Ala Ala Lys Pro Leu Gly Arg Pro Tyr Pro Pro Tyr
                 140                 145                 150
Ser Leu Ala Asp Phe Ser Trp Asn Asn Ile Thr Asp Ser Leu Asp
                 155                 160                 165
Pro Ala Thr Leu Ser Ala Thr Phe Gln Gly His Pro Met Asn Asp
                 170                 175                 180
Pro Thr Arg Thr Phe Ala Asn Gly Ser Leu Ala Phe Arg Val Gln
                 185                 190                 195
Ala Phe Ser Arg Ser Ser Arg Pro Ala Gln Pro Pro Arg Leu Leu
                 200                 205                 210
His Thr Ala Asp Thr Cys Gln Leu Glu Val Ala Leu Ile Gly Ala
                 215                 220                 225
Ser Pro Arg Gly Asn Arg Ser Leu Phe Gly Leu Glu Val Ala Thr
                 230                 235                 240
Leu Gly Gln Gly Pro Asp Cys Pro Ser Met Gln Glu Gln His Ser
                 245                 250                 255
Ile Asp Asp Glu Tyr Ala Pro Ala Val Phe Gln Leu Asp Gln Leu
                 260                 265                 270
Leu Trp Gly Ser Leu Pro Ser Gly Phe Ala Gln Trp Arg Pro Val
                 275                 280                 285
Ala Tyr Ser Gln Lys Pro Gly Gly Arg Glu Ser Ala Leu Pro Cys
                 290                 295                 300
Gln Ala Ser Pro Leu His Pro Ala Leu Ala Tyr Ser Leu Pro Gln
                 305                 310                 315
```

```
Ser Pro Ile Val Arg Ala Phe Phe Gly Ser Gln Asn Asn Phe Cys
                320                 325                 330
Ala Phe Asn Leu Thr Phe Gly Ala Ser Thr Gly Pro Gly Tyr Trp
                335                 340                 345
Asp Gln His Tyr Leu Ser Trp Ser Met Leu Leu Gly Val Gly Phe
                350                 355                 360
Pro Pro Val Asp Gly Leu Ser Pro Leu Val Leu Gly Ile Met Ala
                365                 370                 375
Val Ala Leu Gly Ala Pro Gly Leu Met Leu Leu Gly Gly Gly Leu
                380                 385                 390
Val Leu Leu Leu His His Lys Lys Tyr Ser Glu Tyr Gln Ser Ile
                395                 400                 405
Asn
```

<210> 5
<211> 477
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4620890CD1

<400> 5

```
Met Pro Gly Ser Asp Thr Ala Leu Thr Val Asp Arg Thr Tyr Ser
  1               5                  10                  15
Asp Pro Gly Arg His His Arg Cys Lys Ser Arg Val Glu Arg His
                20                  25                  30
Asp Met Asn Thr Leu Ser Leu Pro Leu Asn Ile Arg Arg Gly Gly
                35                  40                  45
Ser Asp Thr Asn Leu Asn Phe Asp Val Pro Asp Gly Ile Leu Asp
                50                  55                  60
Phe His Lys Val Lys Leu Thr Ala Asp Ser Leu Lys Gln Lys Ile
                65                  70                  75
Leu Lys Val Thr Glu Gln Ile Lys Ile Glu Gln Thr Ser Arg Asp
                80                  85                  90
Gly Asn Val Ala Glu Tyr Leu Lys Leu Val Asn Asn Ala Asp Lys
                95                 100                 105
Gln Gln Ala Gly Arg Ile Lys Gln Val Phe Glu Lys Lys Asn Gln
               110                 115                 120
Lys Ser Ala His Ser Ile Ala Gln Leu Gln Lys Lys Leu Glu Gln
               125                 130                 135
Tyr His Arg Lys Leu Arg Glu Ile Glu Gln Asn Gly Ala Ser Arg
               140                 145                 150
Ser Ser Lys Asp Ile Ser Lys Asp His Leu Lys Asp Ile His Arg
               155                 160                 165
Ser Leu Lys Asp Ala His Val Lys Ser Arg Thr Ala Pro His Cys
               170                 175                 180
Met Glu Ser Ser Lys Ser Gly Met Pro Gly Val Ser Leu Thr Pro
               185                 190                 195
Pro Val Phe Val Phe Asn Lys Ser Arg Glu Phe Ala Asn Leu Ile
               200                 205                 210
Arg Asn Lys Phe Gly Ser Ala Asp Asn Ile Ala His Leu Lys Asn
               215                 220                 225
Ser Leu Glu Glu Phe Arg Pro Glu Ala Ser Ala Arg Ala Tyr Gly
               230                 235                 240
Gly Ser Ala Thr Ile Val Asn Lys Pro Lys Tyr Gly Ser Asp Asp
               245                 250                 255
Glu Cys Ser Ser Gly Thr Ser Gly Ser Ala Asp Ser Asn Gly Asn
               260                 265                 270
Gln Ser Phe Gly Ala Gly Gly Ala Ser Thr Leu Asp Ser Gln Gly
               275                 280                 285
```

```
Lys Leu Ala Val Ile Leu Glu Glu Leu Arg Glu Ile Lys Asp Thr
                290                 295                 300
Gln Ala Gln Leu Ala Glu Asp Ile Glu Ala Leu Lys Val Gln Phe
                305                 310                 315
Lys Arg Glu Tyr Gly Phe Ile Ser Gln Thr Leu Gln Glu Glu Arg
                320                 325                 330
Tyr Arg Tyr Glu Arg Leu Glu Asp Gln Leu His Asp Leu Thr Asp
                335                 340                 345
Leu His Gln His Glu Thr Ala Asn Leu Lys Gln Glu Leu Ala Ser
                350                 355                 360
Ile Glu Glu Lys Val Ala Tyr Gln Ala Tyr Glu Arg Ser Arg Asp
                365                 370                 375
Ile Gln Glu Ala Leu Glu Ser Cys Gln Thr Arg Ile Ser Lys Leu
                380                 385                 390
Glu Leu His Gln Gln Glu Gln Gln Ala Leu Gln Thr Asp Thr Val
                395                 400                 405
Asn Ala Lys Val Leu Leu Gly Arg Cys Ile Asn Val Ile Leu Ala
                410                 415                 420
Phe Met Thr Val Ile Leu Val Cys Val Ser Thr Ile Ala Lys Phe
                425                 430                 435
Val Ser Pro Met Met Lys Ser Arg Cys His Ile Leu Gly Thr Phe
                440                 445                 450
Phe Ala Val Thr Leu Leu Ala Ile Phe Cys Lys Asn Trp Asp His
                455                 460                 465
Ile Leu Cys Ala Ile Glu Arg Met Ile Ile Pro Arg
                470                 475
```

<210> 6
<211> 691
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5514146CD1

<400> 6

```
Met Cys Phe Arg Thr Lys Leu Ser Val Ser Trp Val Pro Leu Phe
  1                 5                  10                  15
Leu Leu Leu Ser Arg Val Phe Ser Thr Glu Thr Asp Lys Pro Ser
                 20                  25                  30
Ala Gln Asp Ser Arg Ser Arg Gly Ser Ser Gly Gln Pro Ala Asp
                 35                  40                  45
Leu Leu Gln Val Leu Ser Ala Gly Asp His Pro Pro His Asn His
                 50                  55                  60
Ser Arg Ser Leu Ile Lys Thr Leu Leu Glu Lys Thr Gly Cys Pro
                 65                  70                  75
Arg Arg Arg Asn Gly Met Gln Gly Asp Cys Asn Leu Cys Phe Glu
                 80                  85                  90
Pro Asp Ala Leu Leu Leu Ile Ala Gly Gly Asn Phe Glu Asp Gln
                 95                 100                 105
Leu Arg Glu Glu Val Val Gln Arg Val Ser Leu Leu Leu Leu Tyr
                110                 115                 120
Tyr Ile Ile His Gln Glu Glu Ile Cys Ser Ser Lys Leu Asn Met
                125                 130                 135
Ser Asn Lys Glu Tyr Lys Phe Tyr Leu His Ser Leu Leu Ser Leu
                140                 145                 150
Arg Gln Asp Glu Asp Ser Ser Phe Leu Ser Gln Asn Glu Thr Glu
                155                 160                 165
Asp Ile Leu Ala Phe Thr Arg Gln Tyr Phe Asp Thr Ser Gln Ser
                170                 175                 180
Gln Cys Met Glu Thr Lys Thr Leu Gln Lys Lys Ser Gly Ile Val
                185                 190                 195
```

```
Ser Ser Glu Gly Ala Asn Glu Ser Thr Leu Pro Gln Leu Ala Ala
                200                 205                 210
Met Ile Ile Thr Leu Ser Leu Gln Gly Val Cys Leu Gly Gln Gly
                215                 220                 225
Asn Leu Pro Ser Pro Asp Tyr Phe Thr Glu Tyr Ile Phe Ser Ser
                230                 235                 240
Leu Asn Arg Thr Asn Thr Leu Arg Leu Ser Glu Leu Asp Gln Leu
                245                 250                 255
Leu Asn Thr Leu Trp Thr Arg Ser Thr Cys Ile Lys Asn Glu Lys
                260                 265                 270
Ile His Gln Phe Gln Arg Lys Gln Asn Asn Ile Ile Thr His Asp
                275                 280                 285
Gln Asp Tyr Ser Asn Phe Ser Ser Ser Met Glu Lys Glu Ser Glu
                290                 295                 300
Asp Gly Pro Ile Ser Trp Asp Gln Thr Cys Phe Ser Ala Arg Gln
                305                 310                 315
Leu Val Glu Ile Phe Leu Gln Lys Gly Leu Ser Leu Ile Ser Lys
                320                 325                 330
Glu Asp Phe Lys Gln Met Ser Pro Gly Ile Ile Gln Gln Leu Leu
                335                 340                 345
Ser Cys Ser Cys His Leu Pro Lys Asp Gln Gln Ala Lys Leu Pro
                350                 355                 360
Pro Thr Thr Leu Glu Lys Tyr Gly Tyr Ser Thr Val Ala Val Thr
                365                 370                 375
Leu Leu Thr Leu Gly Ser Met Leu Gly Thr Ala Leu Val Leu Phe
                380                 385                 390
His Ser Cys Glu Glu Asn Tyr Arg Leu Ile Leu Gln Leu Phe Val
                395                 400                 405
Gly Leu Ala Val Gly Thr Leu Ser Gly Asp Ala Leu Leu His Leu
                410                 415                 420
Ile Pro Gln Val Leu Gly Leu His Lys Gln Glu Ala Pro Glu Phe
                425                 430                 435
Gly His Phe His Glu Ser Lys Gly His Ile Trp Lys Leu Met Gly
                440                 445                 450
Leu Ile Gly Gly Ile His Gly Phe Phe Leu Ile Glu Lys Cys Phe
                455                 460                 465
Ile Leu Leu Val Ser Pro Asn Asp Lys Gln Gly Leu Ser Leu Val
                470                 475                 480
Asn Gly His Val Gly His Ser His His Leu Ala Leu Asn Ser Glu
                485                 490                 495
Leu Ser Asp Gln Ala Gly Arg Gly Lys Ser Ala Ser Thr Ile Gln
                500                 505                 510
Leu Lys Ser Pro Glu Asp Ser Gln Ala Ala Glu Met Pro Ile Gly
                515                 520                 525
Ser Met Thr Ala Ser Asn Arg Lys Cys Lys Ala Ile Ser Leu Leu
                530                 535                 540
Ala Ile Met Ile Leu Val Gly Asp Ser Leu His Asn Phe Ala Asp
                545                 550                 555
Gly Leu Ala Ile Gly Ala Ala Phe Ser Ser Ser Ser Glu Ser Gly
                560                 565                 570
Val Thr Thr Thr Ile Ala Ile Leu Cys His Glu Ile Pro His Glu
                575                 580                 585
Met Gly Asp Phe Ala Val Leu Leu Ser Ser Gly Leu Ser Met Lys
                590                 595                 600
Thr Ala Ile Leu Met Asn Phe Ile Ser Ser Leu Thr Ala Phe Met
                605                 610                 615
Gly Leu Tyr Ile Gly Leu Ser Val Ser Ala Asp Pro Cys Val Gln
                620                 625                 630
Asp Trp Ile Phe Thr Val Thr Ala Gly Met Phe Leu Tyr Leu Ser
                635                 640                 645
Leu Val Glu Met Leu Pro Glu Met Thr His Val Gln Thr Gln Arg
                650                 655                 660
Pro Trp Met Met Phe Leu Leu Gln Asn Phe Gly Leu Ile Leu Gly
```

```
                665                 670                 675
Trp Leu Ser Leu Leu Leu Leu Ala Ile Tyr Glu Gln Asn Ile Lys
                680                 685                 690
Ile


<210> 7
<211> 919
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7474769CD1

<400> 7
Met Gln Glu Cys Leu Thr Leu Trp Val Phe Ser Pro Leu Ala Leu
  1               5                  10                  15
Thr Asp Ser Gly Tyr Thr Lys Thr Tyr Gln Ala His Ala Lys Gln
                 20                  25                  30
Lys Phe Ser Arg Leu Trp Ser Ser Lys Ser Val Thr Glu Ile His
                 35                  40                  45
Leu Tyr Phe Glu Glu Glu Val Lys Gln Glu Glu Cys Asp His Leu
                 50                  55                  60
Asp Arg Leu Phe Ala Pro Lys Glu Ala Gly Lys Gln Pro Arg Thr
                 65                  70                  75
Val Ile Ile Gln Gly Pro Gln Gly Ile Gly Lys Thr Thr Leu Leu
                 80                  85                  90
Met Lys Leu Met Met Ala Trp Ser Asp Asn Lys Ile Phe Arg Asp
                 95                 100                 105
Arg Phe Leu Tyr Thr Phe Tyr Phe Cys Cys Arg Glu Leu Arg Glu
                110                 115                 120
Leu Pro Pro Thr Ser Leu Ala Asp Leu Ile Ser Arg Glu Trp Pro
                125                 130                 135
Asp Pro Ala Ala Pro Ile Thr Glu Ile Val Ser Gln Pro Glu Arg
                140                 145                 150
Leu Leu Phe Val Ile Asp Ser Phe Glu Glu Leu Gln Gly Gly Leu
                155                 160                 165
Asn Glu Pro Asp Ser Asp Leu Cys Gly Asp Leu Met Glu Lys Arg
                170                 175                 180
Pro Val Gln Val Leu Leu Ser Ser Leu Leu Arg Lys Lys Met Leu
                185                 190                 195
Pro Glu Ala Ser Leu Leu Ile Ala Ile Lys Pro Val Cys Pro Lys
                200                 205                 210
Glu Leu Arg Asp Gln Val Thr Ile Ser Glu Ile Tyr Gln Pro Arg
                215                 220                 225
Gly Phe Asn Glu Ser Asp Arg Leu Val Tyr Phe Cys Cys Phe Phe
                230                 235                 240
Lys Asp Pro Lys Arg Ala Met Glu Ala Phe Asn Leu Val Arg Glu
                245                 250                 255
Ser Glu Gln Leu Phe Ser Ile Cys Gln Ile Pro Leu Leu Cys Trp
                260                 265                 270
Ile Leu Cys Thr Ser Leu Lys Gln Glu Met Gln Lys Gly Lys Asp
                275                 280                 285
Leu Ala Leu Thr Cys Gln Ser Thr Thr Ser Val Tyr Ser Ser Phe
                290                 295                 300
Val Phe Asn Leu Phe Thr Pro Glu Gly Ala Glu Gly Pro Thr Pro
                305                 310                 315
Gln Thr Gln His Gln Leu Lys Ala Leu Cys Ser Leu Ala Ala Glu
                320                 325                 330
Gly Met Trp Thr Asp Thr Phe Glu Phe Cys Glu Asp Asp Leu Arg
                335                 340                 345
Arg Asn Gly Val Val Asp Ala Asp Ile Pro Ala Leu Leu Gly Thr
```

```
                350                 355                 360
Lys Ile Leu Leu Lys Tyr Gly Glu Arg Glu Ser Ser Tyr Val Phe
                365                 370                 375
Leu His Val Cys Ile Gln Glu Phe Cys Ala Ala Leu Phe Tyr Leu
                380                 385                 390
Leu Lys Ser His Leu Asp His Pro His Pro Ala Val Arg Cys Val
                395                 400                 405
Gln Glu Leu Leu Val Ala Asn Phe Glu Lys Ala Arg Arg Ala His
                410                 415                 420
Trp Ile Phe Leu Gly Cys Phe Leu Thr Gly Leu Leu Asn Lys Lys
                425                 430                 435
Glu Gln Glu Lys Leu Asp Ala Phe Phe Gly Phe Gln Leu Ser Gln
                440                 445                 450
Glu Ile Lys Gln Gln Ile His Gln Cys Leu Lys Ser Leu Gly Glu
                455                 460                 465
Arg Gly Asn Pro Gln Gly Gln Val Asp Ser Leu Ala Ile Phe Tyr
                470                 475                 480
Cys Leu Phe Glu Met Gln Asp Pro Ala Phe Val Lys Gln Ala Val
                485                 490                 495
Asn Leu Leu Gln Glu Ala Asn Phe His Ile Ile Asp Asn Val Asp
                500                 505                 510
Leu Val Val Ser Ala Tyr Cys Leu Lys Tyr Cys Ser Ser Leu Arg
                515                 520                 525
Lys Leu Cys Phe Ser Val Gln Asn Val Phe Lys Lys Glu Asp Glu
                530                 535                 540
His Ser Ser Thr Ser Asp Tyr Ser Leu Ile Cys Trp His His Ile
                545                 550                 555
Cys Ser Val Leu Thr Thr Ser Gly His Leu Arg Glu Leu Gln Val
                560                 565                 570
Gln Asp Ser Thr Leu Ser Glu Ser Thr Phe Val Thr Trp Cys Asn
                575                 580                 585
Gln Leu Arg His Pro Ser Cys Arg Leu Gln Lys Leu Gly Ile Asn
                590                 595                 600
Asn Val Ser Phe Ser Gly Gln Ser Val Leu Leu Phe Glu Val Leu
                605                 610                 615
Phe Tyr Gln Pro Asp Leu Lys Tyr Leu Ser Phe Thr Leu Thr Lys
                620                 625                 630
Leu Ser Arg Asp Asp Ile Arg Ser Leu Cys Asp Ala Leu Asn Tyr
                635                 640                 645
Pro Ala Gly Asn Val Lys Glu Leu Ala Leu Val Asn Cys His Leu
                650                 655                 660
Ser Pro Ile Asp Cys Glu Val Leu Ala Gly Leu Leu Thr Asn Asn
                665                 670                 675
Lys Lys Leu Thr Tyr Leu Asn Val Ser Cys Asn Gln Leu Asp Thr
                680                 685                 690
Gly Val Pro Leu Leu Cys Glu Ala Leu Cys Ser Pro Asp Thr Val
                695                 700                 705
Leu Val Tyr Leu Met Leu Ala Phe Cys His Leu Ser Glu Gln Cys
                710                 715                 720
Cys Glu Tyr Ile Ser Glu Met Leu Leu Arg Asn Lys Ser Val Arg
                725                 730                 735
Tyr Leu Asp Leu Ser Ala Asn Val Leu Lys Asp Glu Gly Leu Lys
                740                 745                 750
Thr Leu Cys Glu Ala Leu Lys His Pro Asp Cys Cys Leu Asp Ser
                755                 760                 765
Leu Cys Leu Val Lys Cys Phe Ile Thr Ala Ala Gly Cys Glu Asp
                770                 775                 780
Leu Ala Ser Ala Leu Ile Ser Asn Gln Asn Leu Lys Ile Leu Gln
                785                 790                 795
Ile Gly Cys Asn Glu Ile Gly Asp Val Gly Val Gln Leu Leu Cys
                800                 805                 810
Arg Ala Leu Thr His Thr Asp Cys Arg Leu Glu Ile Leu Gly Leu
                815                 820                 825
```

```
Glu Glu Cys Gly Leu Thr Ser Thr Cys Cys Lys Asp Leu Ala Ser
                830                 835                 840
Val Leu Thr Cys Ser Lys Thr Leu Gln Gln Leu Asn Leu Thr Leu
                845                 850                 855
Asn Thr Leu Asp His Thr Gly Val Val Val Leu Cys Glu Ala Leu
                860                 865                 870
Arg His Pro Glu Cys Ala Leu Gln Val Leu Gly Leu Arg Lys Thr
                875                 880                 885
Asp Phe Asp Glu Glu Thr Gln Ala Leu Leu Thr Ala Glu Glu Glu
                890                 895                 900
Arg Asn Pro Asn Leu Thr Ile Thr Asp Asp Cys Asp Thr Ile Thr
                905                 910                 915
Arg Val Glu Ile
```

<210> 8
<211> 178
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 065296CD1

<400> 8

```
Met Ala Gly Lys Pro Tyr Ser Leu Arg Gly Ser Ser His Thr Thr
  1                 5                  10                  15
Gly Thr Phe Leu Leu Leu Ser Gln Ser Ser Gly Glu Leu Gln Ile
                 20                  25                  30
Ile Lys Tyr Phe Lys Met Lys Phe Lys Thr Glu Met Phe Leu Leu
                 35                  40                  45
Leu Leu Leu Trp Arg Asp Cys Met Lys Thr His Thr Gly Met Asn
                 50                  55                  60
His Arg Leu His Val Pro Glu Leu Ser Asn Ala Gln Asp Asn Asn
                 65                  70                  75
Ser Ser Ala Ser Ile Ser Asp Lys Val Gly Phe Ser Lys Ala Glu
                 80                  85                  90
Leu Arg Met Cys Leu Ala Ile Trp Thr Phe Ser Pro Ile Lys Gln
                 95                 100                 105
Val Tyr Lys Ile Leu Lys Ile Glu Cys Leu Asn Phe Ser Ile Val
                110                 115                 120
Leu Ser Val Leu Lys Pro Ile Arg Ile Pro Arg Ile Asn Met Phe
                125                 130                 135
Val Phe Leu Gly Ala Leu Ser Met Thr Gln Asp Asn Glu Trp Tyr
                140                 145                 150
Leu Asn Tyr Ile Phe Phe Thr Leu Glu Ile Ser Arg Gln Lys Val
                155                 160                 165
Phe Phe Glu Trp Val Asn Ser Ala Leu Ser Phe Ser Gln
                170                 175
```

<210> 9
<211> 310
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 231994CD1

<400> 9

```
Met Ser Cys Pro Val Gln Thr Met Asp Pro Glu Val Thr Leu Leu
  1                 5                  10                  15
Leu Gln Cys Pro Gly Gly Gly Leu Pro Gln Glu Gln Ile Gln Ala
```

```
                  20                  25                  30
Glu Leu Ser Pro Ala His Asp Arg Arg Pro Leu Pro Gly Gly Asp
                  35                  40                  45
Glu Ala Ile Thr Ala Ile Trp Glu Thr Arg Leu Lys Ala Gln Pro
                  50                  55                  60
Trp Leu Phe Asp Ala Pro Lys Phe Arg Leu His Ser Ala Thr Leu
                  65                  70                  75
Ala Pro Ile Gly Ser Arg Gly Pro Gln Leu Leu Leu Arg Leu Gly
                  80                  85                  90
Leu Thr Ser Tyr Arg Asp Phe Leu Gly Thr Asn Trp Ser Ser Ser
                  95                 100                 105
Ala Ala Trp Leu Arg Gln Gln Gly Ala Thr Asp Trp Gly Asp Thr
                 110                 115                 120
Gln Ala Tyr Leu Ala Asp Pro Leu Gly Val Gly Ala Ala Leu Ala
                 125                 130                 135
Thr Ala Asp Asp Phe Leu Val Phe Leu Arg Arg Ser Arg Gln Val
                 140                 145                 150
Ala Glu Ala Pro Gly Leu Val Asp Val Pro Gly Gly His Pro Glu
                 155                 160                 165
Pro Gln Ala Leu Cys Pro Gly Gly Ser Pro Gln His Gln Asp Leu
                 170                 175                 180
Ala Gly Gln Leu Val Val His Glu Leu Phe Ser Ser Val Leu Gln
                 185                 190                 195
Glu Ile Cys Asp Glu Val Asn Leu Pro Leu Leu Thr Leu Ser Gln
                 200                 205                 210
Pro Leu Leu Leu Gly Ile Ala Arg Asn Glu Thr Ser Ala Gly Arg
                 215                 220                 225
Ala Ser Ala Glu Phe Tyr Val Gln Cys Ser Leu Thr Ser Glu Gln
                 230                 235                 240
Val Arg Lys His Tyr Leu Ser Gly Gly Pro Glu Ala His Glu Ser
                 245                 250                 255
Thr Gly Ile Phe Phe Val Glu Thr Gln Asn Val Arg Arg Leu Pro
                 260                 265                 270
Glu Thr Glu Met Trp Ala Glu Leu Cys Pro Ser Ala Lys Gly Ala
                 275                 280                 285
Ile Ile Leu Tyr Asn Arg Val Gln Gly Ser Pro Thr Gly Ala Ala
                 290                 295                 300
Leu Gly Ser Pro Ala Leu Leu Pro Pro Leu
                 305                 310
```

<210> 10
<211> 559
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 538054CD1

<400> 10

```
Met Trp Thr Asn Phe Phe Lys Leu Arg Leu Phe Cys Cys Leu Leu
  1               5                  10                  15
Ala Val Leu Met Val Val Val Leu Val Ile Asn Val Thr Gln Val
                 20                  25                  30
Glu Tyr Leu Asp His Glu Thr Val Ser Ala Thr Phe Ile Asp Ser
                 35                  40                  45
Ser Gly Gln Phe Val Ser Ser Gln Val Thr Gly Ile Ser Arg Asn
                 50                  55                  60
Pro Tyr Cys Gly Tyr Asp Gln Gln Thr Leu Ser Ser Gln Glu Arg
                 65                  70                  75
Met Glu Glu Asp Ser Leu Leu Ala Ala Leu His Arg Gln Val Pro
                 80                  85                  90
Asp Val Gly Pro Val Pro Phe Val Lys Ser Thr Asp Pro Ser Ser
```

```
                95                  100                 105
Ser Tyr Phe Val Ile Leu Asn Ser Ala Ala Phe Phe Lys Val Gly
               110                 115                 120
Ser Gln Leu Glu Val Leu Val His Val Gln Asp Phe Gln Arg Lys
               125                 130                 135
Pro Lys Lys Tyr Gly Gly Asp Tyr Leu Gln Ala Arg Ile His Ser
               140                 145                 150
Leu Lys Leu Gln Ala Gly Ala Val Gly Arg Val Val Asp Tyr Gln
               155                 160                 165
Asn Gly Phe Tyr Lys Val Phe Phe Thr Leu Leu Trp Pro Gly Lys
               170                 175                 180
Val Lys Val Ser Val Ser Leu Val His Pro Ser Glu Gly Ile Arg
               185                 190                 195
Val Leu Gln Arg Leu Gln Glu Asp Lys Pro Asp Arg Val Tyr Phe
               200                 205                 210
Lys Ser Leu Phe Arg Ser Gly Arg Ile Ser Glu Thr Thr Glu Cys
               215                 220                 225
Asn Val Cys Leu Pro Gly Asn Leu Pro Leu Cys Asn Phe Thr Asp
               230                 235                 240
Leu Tyr Thr Gly Glu Pro Trp Phe Cys Phe Lys Pro Lys Lys Leu
               245                 250                 255
Pro Cys Ser Ser Arg Ile Thr His Phe Lys Gly Gly Tyr Leu Lys
               260                 265                 270
Gly Leu Leu Thr Ala Ala Glu Ser Ala Phe Phe Gln Ser Gly Val
               275                 280                 285
Asn Ile Lys Met Pro Val Asn Ser Ser Gly Pro Asp Trp Val Thr
               290                 295                 300
Val Ile Pro Arg Arg Ile Lys Glu Thr Asn Ser Leu Glu Leu Ser
               305                 310                 315
Gln Gly Ser Gly Thr Phe Pro Ser Gly Tyr Tyr Tyr Lys Asp Gln
               320                 325                 330
Trp Arg Pro Arg Lys Phe Lys Met Arg Gln Phe Asn Asp Pro Asp
               335                 340                 345
Asn Ile Thr Glu Cys Leu Gln Arg Lys Val Val His Leu Phe Gly
               350                 355                 360
Asp Ser Thr Ile Arg Gln Trp Phe Glu Tyr Leu Thr Thr Phe Val
               365                 370                 375
Pro Asp Leu Val Glu Phe Asn Leu Gly Ser Pro Lys Asn Val Gly
               380                 385                 390
Pro Phe Leu Ala Val Asp Gln Lys His Asn Ile Leu Leu Lys Tyr
               395                 400                 405
Arg Cys His Gly Pro Pro Ile Arg Phe Thr Thr Val Phe Ser Asn
               410                 415                 420
Glu Leu His Tyr Val Ala Asn Glu Leu Asn Gly Ile Val Gly Gly
               425                 430                 435
Lys Asn Thr Val Val Ala Ile Ala Val Trp Ser His Phe Ser Thr
               440                 445                 450
Phe Pro Leu Glu Val Tyr Ile Arg Arg Leu Arg Asn Ile Arg Arg
               455                 460                 465
Ala Val Val Arg Leu Leu Asp Arg Ser Pro Lys Thr Val Val Val
               470                 475                 480
Ile Arg Thr Ala Asn Ala Gln Glu Leu Gly Pro Glu Val Ser Leu
               485                 490                 495
Phe Asn Ser Asp Trp Tyr Asn Phe Gln Leu Asp Thr Ile Leu Arg
               500                 505                 510
Arg Met Phe Ser Gly Val Gly Val Tyr Leu Val Asp Ala Trp Glu
               515                 520                 525
Met Thr Leu Ala His Tyr Leu Pro His Lys Leu His Pro Asp Glu
               530                 535                 540
Val Ile Val Lys Asn Gln Leu Asp Met Phe Leu Ser Phe Val Cys
               545                 550                 555
Pro Leu Glu Thr
```

<210> 11
<211> 477
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1259305CD1

<400> 11
Met Val Cys Val Phe Val Met Asn Arg Met Asn Ser Gln Asn Ser
1 5 10 15
Gly Phe Thr Gln Arg Arg Arg Met Ala Leu Gly Ile Val Ile Leu
20 25 30
Leu Leu Val Asp Val Ile Trp Val Ala Ser Ser Glu Leu Thr Ser
35 40 45
Tyr Val Phe Thr Gln Tyr Asn Lys Pro Phe Phe Ser Thr Phe Ala
50 55 60
Lys Thr Ser Met Phe Val Leu Tyr Leu Leu Gly Phe Ile Ile Trp
65 70 75
Lys Pro Trp Arg Gln Gln Cys Thr Arg Gly Leu Arg Gly Lys His
80 85 90
Ala Ala Phe Phe Ala Asp Ala Glu Gly Tyr Phe Ala Ala Cys Thr
95 100 105
Thr Asp Thr Thr Met Asn Ser Ser Leu Ser Glu Pro Leu Tyr Val
110 115 120
Pro Val Lys Phe His Asp Leu Pro Ser Glu Lys Pro Glu Ser Thr
125 130 135
Asn Ile Asp Thr Glu Lys Thr Pro Lys Lys Ser Arg Val Arg Phe
140 145 150
Ser Asn Ile Met Glu Ile Arg Gln Leu Pro Ser Ser His Ala Leu
155 160 165
Glu Ala Lys Leu Ser Arg Met Ser Tyr Pro Val Lys Glu Gln Glu
170 175 180
Ser Ile Leu Lys Thr Val Gly Lys Leu Thr Ala Thr Gln Val Ala
185 190 195
Lys Ile Ser Phe Phe Phe Cys Phe Val Trp Phe Leu Ala Asn Leu
200 205 210
Ser Tyr Gln Glu Ala Leu Ser Asp Thr Gln Val Ala Ile Val Asn
215 220 225
Ile Leu Ser Ser Thr Ser Gly Leu Phe Thr Leu Ile Leu Ala Ala
230 235 240
Val Phe Pro Ser Asn Ser Gly Asp Arg Phe Thr Leu Ser Lys Leu
245 250 255
Leu Ala Val Ile Leu Ser Ile Gly Gly Val Val Leu Val Asn Leu
260 265 270
Ala Gly Ser Glu Lys Pro Ala Gly Arg Asp Thr Val Gly Ser Ile
275 280 285
Trp Ser Leu Ala Gly Ala Met Leu Tyr Ala Val Tyr Ile Val Met
290 295 300
Ile Lys Arg Lys Val Asp Arg Glu Asp Lys Leu Asp Ile Pro Met
305 310 315
Phe Phe Gly Phe Val Gly Leu Phe Asn Leu Leu Leu Leu Trp Pro
320 325 330
Gly Phe Phe Leu Leu His Tyr Thr Gly Phe Glu Asp Phe Glu Phe
335 340 345
Pro Asn Lys Val Val Leu Met Cys Ile Ile Ile Asn Gly Leu Ile
350 355 360
Gly Thr Val Leu Ser Glu Phe Leu Trp Leu Trp Gly Cys Phe Leu
365 370 375
Thr Ser Ser Leu Ile Gly Thr Leu Ala Leu Ser Leu Thr Ile Pro
380 385 390
Leu Ser Ile Ile Ala Asp Met Cys Met Gln Lys Val Gln Phe Ser

```
                395                 400                 405
Trp Leu Phe Phe Ala Gly Ala Ile Pro Val Phe Phe Ser Phe Phe
                410                 415                 420
Ile Val Thr Leu Leu Cys His Tyr Asn Asn Trp Asp Pro Val Met
                425                 430                 435
Val Gly Ile Arg Arg Ile Phe Ala Phe Ile Cys Arg Lys His Arg
                440                 445                 450
Ile Gln Arg Val Pro Glu Asp Ser Glu Gln Cys Glu Ser Leu Ile
                455                 460                 465
Ser Met His Ser Val Ser Gln Glu Asp Gly Ala Ser
                470                 475
```

<210> 12
<211> 176
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1483702CD1

<400> 12

```
Met Leu Ser Leu Lys Leu Pro Gln Leu Leu Gln Val His Gln Val
  1               5                  10                  15
Pro Arg Val Phe Trp Glu Asp Gly Ile Met Ser Gly Tyr Arg Arg
                 20                  25                  30
Pro Thr Ser Ser Ala Leu Asp Cys Val Leu Ser Ser Phe Gln Met
                 35                  40                  45
Thr Asn Glu Thr Val Asn Ile Trp Thr His Phe Leu Pro Thr Trp
                 50                  55                  60
Tyr Phe Leu Trp Arg Leu Leu Ala Leu Ala Gly Gly Pro Gly Phe
                 65                  70                  75
Arg Ala Glu Pro Tyr His Trp Pro Leu Leu Val Phe Leu Leu Pro
                 80                  85                  90
Ala Cys Leu Tyr Pro Phe Ala Ser Cys Cys Ala His Thr Phe Ser
                 95                 100                 105
Ser Met Ser Pro Arg Met Arg His Ile Cys Tyr Phe Leu Asp Tyr
                110                 115                 120
Gly Ala Leu Ser Leu Tyr Ser Leu Val Ser Trp Ser Trp Lys Ala
                125                 130                 135
Leu Gly Ser Val Arg Ser Ser Ala Gln Glu Pro Ser Pro Ile His
                140                 145                 150
Ser Cys Ser Thr Thr Ser His Ser Phe Ile Gly Ser Gly Cys Ala
                155                 160                 165
Gly Ala Gly Ala Thr Ala Val Gly Arg Arg Pro
                170                 175
```

<210> 13
<211> 190
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1519324CD1

<400> 13

```
Met Ala Gly Met Met Lys Gly Ile Arg Trp Ser Cys Pro Ala Ile
  1               5                  10                  15
Ala Ser Ile Ser Gln Thr Arg Ser Ser Gln Glu Lys Asp Ser Ser
                 20                  25                  30
Ser Pro Pro Trp Asp Leu Arg Arg Ala Ala Thr Glu Trp Gly Gly
                 35                  40                  45
```

```
Pro Arg Cys Ala Val Pro Lys Pro Gly Pro Arg Pro Lys Phe Ser
                50                  55                  60
Leu Pro Ser Leu Val Pro Ser Cys Pro Phe Leu Leu His Ala Trp
                65                  70                  75
Ala Cys Arg Pro Thr Pro Ala Thr Thr Glu Ser Thr Arg Ser Ala
                80                  85                  90
Leu Cys Ser Trp Arg Arg His Ser Arg Val Glu Ser Cys Pro Ser
                95                 100                 105
Leu Ser Leu Gly His Leu Gly Gly Glu Ser Gly Leu Arg Ser Glu
               110                 115                 120
Leu Asp Pro Gly Asp Leu Gly Ser Phe Phe Leu Ala His Gln Pro
               125                 130                 135
Cys Arg Pro His Leu Ser Gln Asn Pro Leu Cys Leu Gly Gly Ser
               140                 145                 150
Gly Ser Ala Leu Leu Cys Ser Arg Gly Trp Gly Val Asp Ser Ile
               155                 160                 165
Arg Trp Glu Ser Gly Val His Pro His Val Ser Val Gly Phe Ser
               170                 175                 180
Pro Trp Gly Trp Lys Lys Arg Ala Ser Thr
               185                 190
```

<210> 14
<211> 75
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1630169CD1

<400> 14

```
Met Ala Leu Gly Lys Val Leu Ala Met Ala Leu Val Leu Ala Leu
  1                 5                  10                  15
Ala Val Leu Gly Ser Leu Ser Pro Gly Ala Arg Ala Gly Asp Cys
                20                  25                  30
Lys Gly Gln Arg Gln Val Leu Arg Glu Ala Pro Gly Phe Val Thr
                35                  40                  45
Asp Gly Ala Gly Asn Tyr Ser Val Asn Gly Asn Cys Glu Trp Leu
                50                  55                  60
Ile Glu Gly Glu Trp Gly Arg Val Gly His Ser Leu Ile Arg Trp
                65                  70                  75
```

<210> 15
<211> 265
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1664253CD1

<400> 15

```
Met Ser Ala Ala Pro Leu Val Gly Tyr Ser Ser Ser Gly Ser Glu
  1                 5                  10                  15
Asp Glu Ser Glu Asp Gly Met Arg Thr Arg Pro Gly Asp Gly Ser
                20                  25                  30
His Arg Arg Gly Gln Ser Pro Leu Pro Arg Gln Arg Phe Pro Val
                35                  40                  45
Pro Asp Ser Val Leu Asn Met Phe Pro Gly Thr Glu Glu Gly Pro
                50                  55                  60
Glu Asp Asp Ser Thr Lys His Gly Gly Arg Val Arg Thr Phe Pro
                65                  70                  75
```

```
His Glu Arg Gly Asn Trp Ala Thr His Val Tyr Val Pro Tyr Glu
                80                  85                  90
Ala Lys Glu Glu Phe Leu Asp Leu Leu Asp Val Leu Leu Pro His
                95                 100                 105
Ala Gln Thr Tyr Val Pro Arg Leu Val Arg Met Lys Val Phe His
               110                 115                 120
Leu Ser Leu Ser Gln Ser Val Val Leu Arg His His Trp Ile Leu
               125                 130                 135
Pro Phe Val Gln Ala Leu Lys Ala Arg Met Thr Ser Phe His Arg
               140                 145                 150
Phe Phe Phe Thr Ala Asn Gln Val Lys Ile Tyr Thr Asn Gln Glu
               155                 160                 165
Lys Thr Arg Thr Phe Ile Gly Leu Glu Val Thr Ser Gly His Ala
               170                 175                 180
Gln Phe Leu Asp Leu Val Ser Glu Val Asp Arg Val Met Glu Glu
               185                 190                 195
Phe Asn Leu Thr Thr Phe Tyr Gln Asp Pro Ser Phe His Leu Ser
               200                 205                 210
Leu Ala Trp Cys Val Gly Asp Ala Arg Leu Gln Leu Glu Gly Gln
               215                 220                 225
Cys Leu Gln Glu Leu Gln Ala Ile Val Asp Gly Phe Glu Asp Ala
               230                 235                 240
Glu Val Leu Leu Arg Val His Thr Glu Gln Val Arg Cys Lys Ser
               245                 250                 255
Gly Asn Lys Phe Phe Ser Met Pro Leu Lys
               260                 265
```

<210> 16
<211> 202
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1864715CD1

<400> 16

```
Met Glu Ala Ala Leu Leu Gly Leu Cys Asn Trp Ser Thr Leu Gly
  1               5                  10                  15
Val Cys Ala Ala Leu Lys Leu Pro Gln Ile Ser Ala Val Leu Ala
                20                  25                  30
Ala Arg Ser Ala Arg Gly Leu Ser Leu Pro Ser Leu Leu Leu Glu
                35                  40                  45
Leu Ala Gly Phe Leu Val Phe Leu Arg Tyr Gln Cys Tyr Tyr Gly
                50                  55                  60
Tyr Pro Pro Leu Thr Tyr Leu Glu Tyr Pro Ile Leu Ile Ala Gln
                65                  70                  75
Asp Val Ile Leu Leu Leu Cys Ile Phe His Phe Asn Gly Asn Val
                80                  85                  90
Lys Gln Ala Thr Pro Tyr Ile Ala Val Leu Val Ser Ser Trp Phe
                95                 100                 105
Ile Leu Ala Leu Gln Lys Trp Ile Ile Asp Leu Ala Met Asn Leu
               110                 115                 120
Cys Thr Phe Ile Ser Ala Ala Ser Lys Phe Ala Gln Leu Gln Cys
               125                 130                 135
Leu Trp Lys Thr Arg Asp Ser Gly Thr Val Ser Ala Leu Thr Trp
               140                 145                 150
Ser Leu Ser Ser Tyr Thr Cys Ala Thr Arg Ile Ile Thr Thr Leu
               155                 160                 165
Met Thr Thr Asn Asp Phe Thr Ile Leu Leu Arg Phe Val Ile Met
               170                 175                 180
Leu Ala Leu Asn Ile Trp Val Thr Val Thr Val Leu Arg Tyr Arg
               185                 190                 195
```

Lys Thr Ala Ile Lys Ala Glu
200

<210> 17
<211> 111
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1929395CD1

<400> 17
Met Tyr Pro Thr Ala Pro Glu Leu Leu Val Pro Gln Pro Arg Pro
1 5 10 15
Gln Gly Ser Pro Ala Ser Leu Leu Leu Gly Thr Pro Val Leu Ala
20 25 30
Ala Val Tyr Gly Ala Ser Cys Leu Pro Leu Gly Arg His Pro Cys
35 40 45
Thr Pro Ala Ser Phe Pro Trp Pro Phe Leu Ala Pro Val Leu Leu
50 55 60
Leu Tyr Ile Asp Leu Phe Thr Gln Lys Arg Ala Arg Pro Leu Phe
65 70 75
Ser Ala Thr Ser Pro Val Ser Glu Ile Gln Pro Pro Arg Leu His
80 85 90
Arg Lys Ile Asp Ile Leu Glu Ile Met Lys Ser Asp Ile Phe Ala
95 100 105
Tyr Glu Arg Lys Lys Gly
110

<210> 18
<211> 105
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1987737CD1

<400> 18
Met Cys Val His Arg Cys Glu Cys Val Cys Met Arg Ala Cys Leu
1 5 10 15
Cys Ala Gly Val Cys Met Cys Val Ala Ser Cys Leu Gly Leu Pro
20 25 30
Met Asn Val Val Glu Cys Tyr Thr Trp Arg Val Leu Val Phe His
35 40 45
Gln Phe Gln Asp Glu Glu Leu His Asp Thr Val Asp Leu Glu Thr
50 55 60
Ile Pro Leu Glu Arg Gln Pro Arg Asp Val Gln His Pro Val Ser
65 70 75
Thr Arg Ile Leu Tyr Leu His Val Tyr Phe Val Ala Val Thr Leu
80 85 90
Thr Leu Ile Arg Ile Leu Gln Leu Trp Thr Glu Ala Phe Ser Pro
95 100 105

<210> 19
<211> 717
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 2122866CD1

<400> 19

```
Met Ala Ser Ser Ser Asp Ser Glu Asp Asp Ser Phe Met Ala Val
  1               5                  10                  15
Asp Gln Glu Glu Thr Val Leu Glu Gly Thr Met Asp Gln Asp Glu
                20                  25                  30
Glu Pro His Pro Val Leu Glu Ala Glu Glu Thr Arg His Asn Arg
                35                  40                  45
Ser Met Ser Glu Leu Pro Glu Glu Val Leu Glu Tyr Ile Leu Ser
                50                  55                  60
Phe Leu Ser Pro Tyr Gln Glu His Lys Thr Ala Ala Leu Val Cys
                65                  70                  75
Lys Gln Trp Tyr Arg Leu Ile Lys Gly Val Ala His Gln Cys Tyr
                80                  85                  90
His Gly Phe Met Lys Ala Val Gln Glu Gly Asn Ile Gln Trp Glu
                95                 100                 105
Ser Arg Thr Tyr Pro Tyr Pro Gly Thr Pro Ile Thr Gln Arg Phe
               110                 115                 120
Ser His Ser Ala Cys Tyr Tyr Asp Ala Asn Gln Ser Met Tyr Val
               125                 130                 135
Phe Gly Gly Cys Thr Gln Ser Ser Cys Asn Ala Ala Phe Asn Asp
               140                 145                 150
Leu Trp Arg Leu Asp Leu Asn Ser Lys Glu Trp Ile Arg Pro Leu
               155                 160                 165
Ala Ser Gly Ser Tyr Pro Ser Pro Lys Ala Gly Ala Thr Leu Val
               170                 175                 180
Val Tyr Lys Asp Leu Leu Val Leu Phe Gly Gly Trp Thr Arg Pro
               185                 190                 195
Ser Pro Tyr Pro Leu His Gln Pro Glu Arg Phe Phe Asp Glu Ile
               200                 205                 210
His Thr Tyr Ser Pro Ser Lys Asn Trp Trp Asn Cys Ile Val Thr
               215                 220                 225
Thr His Gly Pro Pro Pro Met Ala Gly His Ser Ser Cys Val Ile
               230                 235                 240
Asp Asp Lys Met Ile Val Phe Gly Gly Ser Leu Gly Ser Arg Gln
               245                 250                 255
Met Ser Asn Asp Val Trp Val Leu Asp Leu Glu Gln Trp Ala Trp
               260                 265                 270
Ser Lys Pro Asn Ile Ser Gly Pro Ser Pro His Pro Arg Gly Gly
               275                 280                 285
Gln Ser Gln Ile Val Ile Asp Asp Ala Thr Ile Leu Ile Leu Gly
               290                 295                 300
Gly Cys Gly Gly Pro Asn Ala Leu Phe Lys Asp Ala Trp Leu Leu
               305                 310                 315
His Met His Ser Gly Pro Trp Ala Trp Gln Pro Leu Lys Val Glu
               320                 325                 330
Asn Glu Glu His Gly Ala Pro Glu Leu Trp Cys His Pro Ala Cys
               335                 340                 345
Arg Val Gly Gln Cys Val Val Val Phe Ser Gln Ala Pro Ser Gly
               350                 355                 360
Arg Ala Pro Leu Ser Pro Ser Leu Asn Ser Arg Pro Ser Pro Ile
               365                 370                 375
Ser Ala Thr Pro Pro Ala Leu Val Pro Glu Thr Arg Glu Tyr Arg
               380                 385                 390
Ser Gln Ser Pro Val Arg Ser Met Asp Glu Ala Pro Cys Val Asn
               395                 400                 405
Gly Arg Trp Gly Thr Leu Arg Pro Arg Ala Gln Arg Gln Thr Pro
               410                 415                 420
Ser Gly Ser Arg Glu Gly Ser Leu Ser Pro Ala Arg Gly Asp Gly
               425                 430                 435
Ser Pro Ile Leu Asn Gly Gly Ser Leu Ser Pro Gly Thr Ala Ala
               440                 445                 450
```

```
Val Gly Gly Ser Ser Leu Asp Ser Pro Val Gln Ala Ile Ser Pro
                455                 460                 465
Ser Thr Pro Ser Ala Pro Glu Gly Tyr Asp Leu Lys Ile Gly Leu
                470                 475                 480
Ser Leu Ala Pro Arg Arg Gly Ser Leu Pro Asp Gln Lys Asp Leu
                485                 490                 495
Arg Leu Gly Ser Ile Asp Leu Asn Trp Asp Leu Lys Pro Ala Ser
                500                 505                 510
Ser Ser Asn Pro Met Asp Gly Met Asp Asn Arg Thr Val Gly Gly
                515                 520                 525
Ser Met Arg His Pro Pro Glu Gln Thr Asn Gly Val His Thr Pro
                530                 535                 540
Pro His Val Ala Ser Ala Leu Ala Gly Ala Val Ser Pro Gly Ala
                545                 550                 555
Leu Arg Arg Ser Leu Glu Ala Ile Lys Ala Met Ser Ser Lys Gly
                560                 565                 570
Pro Ser Ala Ser Ala Ala Leu Ser Pro Pro Leu Gly Ser Ser Pro
                575                 580                 585
Gly Ser Pro Gly Ser Gln Ser Leu Ser Ser Gly Glu Thr Val Pro
                590                 595                 600
Ile Pro Arg Pro Gly Pro Ala Gln Gly Asp Gly His Ser Leu Pro
                605                 610                 615
Pro Ile Ala Arg Arg Leu Gly His His Pro Pro Gln Ser Leu Asn
                620                 625                 630
Val Gly Lys Pro Leu Tyr Gln Ser Met Asn Cys Lys Pro Met Gln
                635                 640                 645
Met Tyr Val Leu Asp Ile Lys Asp Thr Lys Glu Lys Gly Arg Val
                650                 655                 660
Lys Trp Lys Val Phe Asn Ser Ser Ser Val Val Gly Pro Pro Glu
                665                 670                 675
Thr Ser Leu His Thr Val Val Gln Gly Arg Gly Glu Leu Ile Ile
                680                 685                 690
Phe Gly Gly Leu Met Asp Lys Lys Gln Asn Val Lys Tyr Tyr Pro
                695                 700                 705
Lys Thr Asn Ala Leu Tyr Phe Val Arg Ala Lys Arg
                710                 715

<210> 20
<211> 580
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2123981CD1

<400> 20
Met His Leu Ser Ala Val Phe Asn Ala Leu Leu Val Ser Val Leu
  1               5                  10                  15
Ala Ala Val Leu Trp Lys His Val Arg Leu Arg Glu His Ala Ala
                20                  25                  30
Thr Leu Glu Glu Glu Leu Ala Leu Ser Arg Gln Ala Thr Glu Pro
                35                  40                  45
Ala Pro Ala Leu Arg Ile Asp Tyr Pro Lys Ala Leu Gln Ile Leu
                50                  55                  60
Met Glu Gly Gly Thr His Met Val Cys Thr Gly Arg Thr His Thr
                65                  70                  75
Asp Arg Ile Cys Arg Phe Lys Trp Leu Cys Tyr Ser Asn Glu Ala
                80                  85                  90
Glu Glu Phe Ile Phe Phe His Gly Asn Thr Ser Val Met Leu Pro
                95                 100                 105
Asn Leu Gly Ser Arg Arg Phe Gln Pro Ala Leu Leu Asp Leu Ser
               110                 115                 120
```

```
Thr Val Glu Asp His Asn Thr Gln Tyr Phe Asn Phe Val Glu Leu
                125                 130                 135
Pro Ala Ala Ala Leu Arg Phe Met Pro Lys Pro Val Phe Val Pro
                140                 145                 150
Asp Val Ala Leu Ile Ala Asn Arg Phe Asn Pro Asp Asn Leu Met
                155                 160                 165
His Val Phe His Asp Asp Leu Leu Pro Leu Phe Tyr Thr Leu Arg
                170                 175                 180
Gln Phe Pro Gly Leu Ala His Glu Ala Arg Leu Phe Phe Met Glu
                185                 190                 195
Gly Trp Gly Glu Gly Ala His Phe Asp Leu Tyr Lys Leu Leu Ser
                200                 205                 210
Pro Lys Gln Pro Leu Leu Arg Ala Gln Leu Lys Thr Leu Gly Arg
                215                 220                 225
Leu Leu Cys Phe Ser His Ala Phe Val Gly Leu Ser Lys Ile Thr
                230                 235                 240
Thr Trp Tyr Gln Tyr Gly Phe Val Gln Pro Gln Gly Pro Lys Ala
                245                 250                 255
Asn Ile Leu Val Ser Gly Asn Glu Ile Arg Gln Phe Ala Arg Phe
                260                 265                 270
Met Thr Glu Lys Leu Asn Val Ser His Thr Gly Val Pro Leu Gly
                275                 280                 285
Glu Glu Tyr Ile Leu Val Phe Ser Arg Thr Gln Asn Arg Leu Ile
                290                 295                 300
Leu Asn Glu Ala Glu Leu Leu Leu Ala Leu Ala Gln Glu Phe Gln
                305                 310                 315
Met Lys Thr Val Thr Val Ser Leu Glu Asp His Thr Phe Ala Asp
                320                 325                 330
Val Val Arg Leu Val Ser Asn Ala Ser Met Leu Val Ser Met His
                335                 340                 345
Gly Ala Gln Leu Val Thr Thr Leu Phe Leu Pro Arg Gly Ala Thr
                350                 355                 360
Val Val Glu Leu Phe Pro Tyr Ala Val Asn Pro Asp His Tyr Thr
                365                 370                 375
Pro Tyr Lys Thr Leu Ala Met Leu Pro Gly Met Asp Leu Gln Tyr
                380                 385                 390
Val Ala Trp Arg Asn Met Met Pro Glu Asn Thr Val Thr His Pro
                395                 400                 405
Glu Arg Pro Trp Asp Gln Gly Gly Ile Thr His Leu Asp Arg Ala
                410                 415                 420
Glu Gln Ala Arg Ile Leu Gln Ser Arg Glu Val Pro Arg His Leu
                425                 430                 435
Cys Cys Arg Asn Pro Glu Trp Leu Phe Arg Ile Tyr Gln Asp Thr
                440                 445                 450
Lys Val Asp Ile Pro Ser Leu Ile Gln Thr Ile Arg Arg Val Val
                455                 460                 465
Lys Gly Arg Pro Gly Pro Arg Lys Gln Lys Trp Thr Val Gly Leu
                470                 475                 480
Tyr Pro Gly Lys Val Arg Glu Ala Arg Cys Gln Ala Ser Val His
                485                 490                 495
Gly Ala Ser Glu Ala Arg Leu Thr Val Ser Trp Gln Ile Pro Trp
                500                 505                 510
Asn Leu Lys Tyr Leu Lys Val Arg Glu Val Lys Tyr Glu Val Trp
                515                 520                 525
Leu Gln Glu Gln Gly Glu Asn Thr Tyr Val Pro Tyr Ile Leu Ala
                530                 535                 540
Leu Gln Asn His Thr Phe Thr Glu Asn Ile Lys Pro Phe Thr Thr
                545                 550                 555
Tyr Leu Val Trp Val Arg Cys Ile Phe Asn Lys Ile Leu Leu Gly
                560                 565                 570
Pro Phe Ala Asp Val Leu Val Cys Asn Thr
                575                 580
```

<210> 21
<211> 172
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2200177CD1

<400> 21
Met Leu Ser Glu Asp Lys Leu Phe Gln Ile Ile His Ser Leu Leu
1 5 10 15
Ile Thr Gln Leu Ala Ser Glu Thr Lys Ile Ser Ala Thr Ile Cys
20 25 30
Leu Pro Leu Leu Phe His Cys Leu Phe Leu Leu Val Leu Ser Phe
35 40 45
Pro Ile Thr Leu Cys Ile Arg His Ser Gly Pro Tyr His Ile Tyr
50 55 60
Pro Leu Leu Gln Val Ser Asn Leu Ile Phe Leu Gln Thr His Phe
65 70 75
Leu Ser Tyr Ile Ala Gly Ile Met Gln Lys Leu Leu Ser Asn Val
80 85 90
Val His Ser Gln Lys Ile His Pro Glu Ile Leu Arg Phe Gly Lys
95 100 105
Val Cys Ala Gln Ser Thr Ile Ser Lys Lys Phe Lys Glu Glu Lys
110 115 120
Tyr Lys Thr Pro His Thr Ile Ser Leu Ile Ser Gln Ile His Glu
125 130 135
Thr Ala Thr Ile Lys Ser Lys Val Phe Arg Lys Leu Ser Thr Tyr
140 145 150
Phe Ser Ile Val Leu Lys Leu Lys Glu Ile Lys Ile Ala Gly Phe
155 160 165
Lys Tyr Leu Trp Ser Ser Asn
170

<210> 22
<211> 256
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2319255CD1

<400> 22
Met Asn Thr Phe Lys Ala Val Gly Lys Ile Arg Gly Lys Pro Leu
1 5 10 15
Pro Leu Leu Leu Phe Phe Glu Ala Leu Phe Ile Thr Ser His Ala
20 25 30
Phe Pro Cys Pro Val Asp Ala Ala Leu Thr Leu Glu Gly Ile Lys
35 40 45
Cys Gly Leu Ser Glu Lys Arg Leu Asp Leu Val Thr Asn Trp Val
50 55 60
Thr Gln Glu Arg Leu Thr Phe Ser Glu Glu Ala Gly Asp Val Ile
65 70 75
Cys Asp Tyr Gly Glu Gln Asp Thr Tyr Asn Lys Ala Lys Cys Leu
80 85 90
Ala Leu Ala Gln Ile Ile Tyr Ser Glu Cys Gly Leu His Lys Lys
95 100 105
Ala Ile Leu Cys Leu Cys Lys Gln Gly Gln Thr His Arg Val Met
110 115 120
Glu Tyr Ile Gln Gln Leu Lys Asp Phe Thr Thr Asp Asp Leu Leu
125 130 135

```
Gln Leu Leu Met Ser Cys Pro Gln Val Glu Leu Ile Gln Cys Leu
                140                 145                 150
Thr Lys Glu Leu Asn Glu Lys Gln Pro Ser Leu Ser Phe Gly Leu
                155                 160                 165
Ala Ile Leu His Leu Phe Ser Ala Asp Met Lys Lys Val Gly Ile
                170                 175                 180
Lys Leu Leu Gln Glu Ile Asn Lys Gly Gly Ile Asp Ala Val Glu
                185                 190                 195
Ser Leu Met Ile Asn Asp Ser Phe Cys Ser Ile Glu Lys Trp Gln
                200                 205                 210
Glu Val Ala Asn Ile Cys Ser Gln Asn Gly Phe Asp Lys Leu Ser
                215                 220                 225
Asn Asp Ile Thr Ser Ile Leu Arg Ser Gln Ala Ala Val Thr Glu
                230                 235                 240
Ile Ser Glu Glu Asp Asp Ala Val Asn Leu Met Glu His Val Phe
                245                 250                 255
Trp
```

<210> 23
<211> 93
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2792452CD1

<400> 23

```
Met Ser Pro Met Trp Ala Pro Pro Trp Leu Pro Leu Leu Leu Ser
  1               5                  10                  15
Lys Ser Glu Pro Thr Gln Ser Pro Ser Pro Arg Arg Pro Leu Pro
                20                  25                  30
Pro Gly Lys Met Thr Leu Gly Gln Gly Ser Leu Leu Met Ser Val
                35                  40                  45
Phe Cys Leu Val Gly Leu Gly Val Pro Leu Pro Leu Ile Arg Arg
                50                  55                  60
Gly Phe Arg Ala Glu Ile Lys Pro Gln Thr Gly Glu Pro Leu Trp
                65                  70                  75
His Met Ala Pro Arg Ala Ser His Ala Ser Gly Phe Ser Pro Cys
                80                  85                  90
Gln Asp Thr
```

<210> 24
<211> 112
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2853088CD1

<400> 24

```
Met Pro Thr Asp Val Pro Lys Ala Arg Leu Glu Leu Thr Ser Leu
  1               5                  10                  15
Leu Leu Leu Leu Leu Phe Leu Arg Trp Ser Leu Ala Leu Leu Pro
                20                  25                  30
Arg Leu Asp Cys Ser Gly Ala Val Leu Ala His Cys Asn Phe Arg
                35                  40                  45
Leu Trp Gly Ser Ser Asp Ser Ser Ala Ser Ala Ser Ser Gln Val
                50                  55                  60
Ala Gly Ser Thr Gly Ala Cys His Gln Ala Arg Ala Lys Glu Arg
```

```
                65                  70                  75
Asp Ser Ile Ser Lys Ile Ile Thr Ile Ile Ile Met Arg Ser Ile
                80                  85                  90
Pro Asp Val Leu Leu Gly Arg Leu Trp Ala Tyr Ser Leu Glu Leu
                95                 100                 105
Arg Arg Asp Ile Lys Ala Ser
               110

<210> 25
<211> 186
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2949004CD1

<400> 25
Met Lys Leu Pro Arg Glu Glu Ala Ser Ala Ser Phe Val Arg Arg
  1               5                  10                  15
Ala Asp Leu Thr Arg Glu Asp Leu Ala Pro Ser Ser Val Asp Ser
                20                  25                  30
Gly Gln Ala Gly Phe Gly Gly Cys Cys Glu Ser Gly Leu Pro Asn
                35                  40                  45
Thr Met Pro Ser Ala Phe Ser Val Ser Ser Phe Pro Val Ser Ile
                50                  55                  60
Pro Ala Val Leu Thr Gln Thr Asp Trp Thr Glu Pro Trp Leu Met
                65                  70                  75
Gly Leu Ala Thr Phe His Ala Leu Cys Val Leu Leu Thr Cys Leu
                80                  85                  90
Ser Ser Arg Ser Tyr Arg Leu Gln Ile Gly His Phe Leu Cys Leu
                95                 100                 105
Val Ile Leu Val Tyr Cys Ala Glu Tyr Ile Asn Glu Ala Ala Ala
               110                 115                 120
Met Asn Trp Arg Leu Phe Ser Lys Tyr Gln Tyr Phe Asp Ser Arg
               125                 130                 135
Gly Met Phe Ile Ser Ile Val Phe Ser Ala Pro Leu Leu Val Asn
               140                 145                 150
Ala Met Ile Ile Val Val Met Trp Val Trp Lys Thr Leu Asn Val
               155                 160                 165
Met Thr Asp Leu Lys Asn Ala Gln Glu Arg Arg Lys Glu Lys Lys
               170                 175                 180
Arg Arg Arg Lys Glu Asp
               185

<210> 26
<211> 487
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3011670CD1

<400> 26
Met Gly Val Ile Gly Ile Gln Leu Val Val Thr Met Val Met Ala
  1               5                  10                  15
Ser Val Met Gln Lys Ile Ile Pro His Tyr Ser Leu Ala Arg Trp
                20                  25                  30
Leu Leu Cys Asn Gly Ser Leu Arg Trp Tyr Gln His Pro Thr Glu
                35                  40                  45
Glu Glu Leu Arg Ile Leu Ala Gly Lys Gln Gln Lys Gly Lys Thr
                50                  55                  60
```

```
Lys Lys Asp Arg Lys Tyr Asn Gly His Ile Glu Ser Lys Pro Leu
                65                  70                  75
Thr Ile Pro Lys Asp Ile Asp Leu His Leu Glu Thr Lys Ser Val
                80                  85                  90
Thr Glu Val Asp Thr Leu Ala Leu His Tyr Phe Pro Glu Tyr Gln
                95                 100                 105
Trp Leu Val Asp Phe Thr Val Ala Ala Thr Val Val Tyr Leu Val
               110                 115                 120
Thr Glu Val Tyr Tyr Asn Phe Met Lys Pro Thr Gln Glu Met Asn
               125                 130                 135
Ile Ser Leu Val Trp Cys Leu Leu Val Leu Ser Phe Ala Ile Lys
               140                 145                 150
Val Leu Phe Ser Leu Thr Thr His Tyr Phe Lys Val Glu Asp Gly
               155                 160                 165
Gly Glu Arg Ser Val Cys Val Thr Phe Gly Phe Phe Phe Phe Val
               170                 175                 180
Lys Ala Met Ala Val Leu Ile Val Thr Glu Asn Tyr Leu Glu Phe
               185                 190                 195
Gly Leu Glu Thr Gly Phe Thr Asn Phe Ser Asp Ser Ala Met Gln
               200                 205                 210
Phe Leu Glu Lys Gln Gly Leu Glu Ser Gln Ser Pro Val Ser Lys
               215                 220                 225
Leu Thr Phe Lys Phe Phe Leu Ala Ile Phe Cys Ser Phe Ile Gly
               230                 235                 240
Ala Phe Leu Thr Phe Pro Gly Leu Arg Leu Ala Gln Met His Leu
               245                 250                 255
Asp Ala Leu Asn Leu Ala Thr Glu Lys Ile Thr Gln Thr Leu Leu
               260                 265                 270
His Ile Asn Phe Leu Ala Pro Leu Phe Met Val Leu Leu Trp Val
               275                 280                 285
Lys Pro Ile Thr Lys Asp Tyr Ile Met Asn Pro Pro Leu Gly Lys
               290                 295                 300
Glu Ser Ile Pro Leu Met Thr Glu Ala Thr Phe Asp Thr Leu Arg
               305                 310                 315
Leu Trp Leu Ile Ile Leu Leu Cys Ala Leu Arg Leu Ala Met Met
               320                 325                 330
Arg Ser His Leu Gln Ala Tyr Leu Asn Leu Ala Gln Lys Cys Val
               335                 340                 345
Asp Gln Met Lys Lys Glu Ala Gly Arg Ile Ser Thr Val Glu Leu
               350                 355                 360
Gln Lys Met Val Ala Arg Val Phe Tyr Tyr Leu Cys Val Ile Ala
               365                 370                 375
Leu Gln Tyr Val Ala Pro Leu Val Met Leu Leu His Thr Thr Leu
               380                 385                 390
Leu Leu Lys Thr Leu Gly Asn His Ser Trp Gly Ile Tyr Pro Glu
               395                 400                 405
Ser Ile Ser Thr Leu Pro Val Asp Asn Ser Leu Leu Ser Asn Ser
               410                 415                 420
Val Tyr Ser Glu Leu Pro Ser Ala Glu Gly Lys Met Lys Val Thr
               425                 430                 435
Val Thr Gln Ile Thr Val Ala Leu Ser Ser Leu Lys Asn Ile Phe
               440                 445                 450
Thr Pro Leu Leu Phe Arg Gly Leu Leu Ser Phe Leu Thr Trp Trp
               455                 460                 465
Ile Ala Ala Cys Leu Phe Ser Thr Ser Leu Phe Gly Leu Phe Tyr
               470                 475                 480
His Gln Tyr Leu Thr Val Ala
               485
```

<210> 27
<211> 350
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3242083CD1

<400> 27
Met Asn Pro Arg Gly Leu Phe Gln Asp Phe Asn Pro Ser Lys Phe
1 5 10 15
Leu Ile Tyr Thr Cys Leu Leu Leu Phe Ser Val Leu Leu Pro Leu
20 25 30
Arg Leu Asp Gly Ile Ile Gln Trp Ser Tyr Trp Ala Val Phe Ala
35 40 45
Pro Ile Trp Leu Trp Lys Leu Leu Val Val Ala Gly Ala Ser Val
50 55 60
Gly Ala Gly Val Trp Ala Arg Asn Pro Arg Tyr Arg Thr Glu Gly
65 70 75
Glu Ala Cys Val Glu Phe Lys Ala Met Leu Ile Ala Val Gly Ile
80 85 90
His Leu Leu Leu Leu Met Phe Glu Val Leu Val Cys Asp Arg Val
95 100 105
Glu Arg Gly Thr His Phe Trp Leu Leu Val Phe Met Pro Leu Phe
110 115 120
Phe Val Ser Pro Val Ser Val Ala Ala Cys Val Trp Gly Phe Arg
125 130 135
His Asp Arg Ser Leu Glu Leu Glu Ile Leu Cys Ser Val Asn Ile
140 145 150
Leu Gln Phe Ile Phe Ile Ala Leu Lys Leu Asp Arg Ile Ile His
155 160 165
Trp Pro Trp Leu Val Val Phe Val Pro Leu Trp Ile Leu Met Ser
170 175 180
Phe Leu Cys Leu Val Val Leu Tyr Tyr Ile Val Trp Ser Leu Leu
185 190 195
Phe Leu Arg Ser Leu Asp Val Val Ala Glu Gln Arg Arg Thr His
200 205 210
Val Thr Met Ala Ile Ser Trp Ile Thr Ile Val Val Pro Leu Leu
215 220 225
Thr Phe Glu Val Leu Leu Val His Arg Leu Asp Gly His Asn Thr
230 235 240
Phe Ser Tyr Val Ser Ile Phe Val Pro Leu Trp Leu Ser Leu Leu
245 250 255
Thr Leu Met Ala Thr Thr Phe Arg Arg Lys Gly Gly Asn His Trp
260 265 270
Trp Phe Gly Ile Arg Arg Asp Phe Cys Gln Phe Leu Leu Glu Ile
275 280 285
Phe Pro Phe Leu Arg Glu Tyr Gly Asn Ile Ser Tyr Asp Leu His
290 295 300
His Glu Asp Ser Glu Asp Ala Glu Glu Thr Ser Val Pro Glu Ala
305 310 315
Pro Lys Ile Ala Pro Ile Phe Gly Lys Lys Ala Arg Val Val Ile
320 325 330
Thr Gln Ser Pro Gly Lys Tyr Val Pro Pro Pro Pro Lys Leu Asn
335 340 345
Ile Asp Met Pro Asp
350

<210> 28
<211> 450
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3363391CD1

<400> 28

```
Met Leu Pro Ser Cys Leu Trp Phe Arg Gly Thr Gly Leu Ile Trp
 1               5                  10                  15
Trp Val Thr Gly Thr Ala Ser Ala Ala Gly Leu Leu Tyr Leu His
            20                  25                  30
Thr Trp Ala Ala Ala Val Ser Gly Cys Val Phe Ala Ile Phe Thr
        35                  40                  45
Ala Ser Met Trp Pro Gln Thr Leu Gly His Leu Ile Asn Ser Gly
    50                  55                  60
Thr Asn Pro Gly Lys Thr Met Thr Ile Ala Met Ile Phe Tyr Leu
65                  70                  75
Leu Glu Ile Phe Phe Cys Ala Trp Cys Thr Ala Phe Lys Phe Val
                80                  85                  90
Pro Gly Gly Val Tyr Ala Arg Glu Arg Ser Asp Val Leu Leu Gly
            95                  100                 105
Thr Met Met Leu Ile Ile Gly Leu Asn Met Leu Phe Gly Pro Lys
        110                 115                 120
Lys Asn Leu Asp Leu Leu Leu Gln Thr Lys Asn Ser Ser Lys Val
    125                 130                 135
Leu Phe Arg Lys Ser Glu Lys Tyr Met Lys Leu Phe Leu Trp Leu
140                 145                 150
Leu Val Gly Val Gly Leu Leu Gly Leu Gly Leu Arg His Lys Ala
155                 160                 165
Tyr Glu Arg Lys Leu Gly Lys Val Ala Pro Thr Lys Glu Val Ser
                170                 175                 180
Ala Ala Ile Trp Pro Phe Arg Phe Gly Tyr Asp Asn Glu Gly Trp
            185                 190                 195
Ser Ser Leu Glu Arg Ser Ala His Leu Leu Asn Glu Thr Gly Ala
        200                 205                 210
Asp Phe Ile Thr Ile Leu Glu Ser Asp Ala Ser Lys Pro Tyr Met
    215                 220                 225
Gly Asn Asn Asp Leu Thr Met Trp Leu Gly Glu Lys Leu Gly Phe
230                 235                 240
Tyr Thr Asp Phe Gly Pro Ser Thr Arg Tyr His Thr Trp Gly Ile
                245                 250                 255
Met Ala Leu Ser Arg Tyr Pro Ile Val Lys Ser Glu His His Leu
            260                 265                 270
Leu Pro Ser Pro Glu Gly Glu Ile Ala Pro Ala Ile Thr Leu Thr
        275                 280                 285
Val Asn Ile Ser Gly Lys Leu Val Asp Phe Val Val Thr His Phe
    290                 295                 300
Gly Asn His Glu Asp Asp Leu Asp Arg Lys Leu Gln Ala Ile Ala
305                 310                 315
Val Ser Lys Leu Leu Lys Ser Ser Ser Asn Gln Val Ile Phe Leu
                320                 325                 330
Gly Tyr Ile Thr Ser Ala Pro Gly Ser Arg Asp Tyr Leu Gln Leu
            335                 340                 345
Thr Glu His Gly Asn Val Lys Asp Ile Asp Ser Thr Asp His Asp
        350                 355                 360
Arg Trp Cys Glu Tyr Ile Met Tyr Arg Gly Leu Ile Arg Leu Gly
    365                 370                 375
Tyr Ala Arg Ile Ser His Ala Glu Leu Ser Asp Ser Glu Ile Gln
380                 385                 390
Met Ala Lys Phe Arg Ile Pro Asp Asp Pro Thr Asn Tyr Arg Asp
395                 400                 405
Asn Gln Lys Val Val Ile Asp His Arg Glu Val Ser Glu Lys Ile
                410                 415                 420
His Phe Asn Pro Arg Phe Gly Ser Tyr Lys Glu Gly His Asn Tyr
            425                 430                 435
Glu Asn Asn His His Phe His Met Asn Thr Pro Lys Tyr Phe Leu
        440                 445                 450
```

<210> 29
<211> 400
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3703614CD1

<400> 29
Met Asn Pro Phe Leu Gly Asn Leu Pro Ser Ala Pro Ala Met Gly
1 5 10 15
Cys Ser Asp Ala Ser Thr Leu Asn Pro Gly Ser Ala Ser His Val
20 25 30
Ser Thr Tyr Thr Glu Asp Ser Gly Ser Ala His Gln Ser Arg Asp
35 40 45
Gln Val Phe Leu Pro Ala Phe Pro Val Gln Val Arg Arg Cys Lys
50 55 60
Ala Leu Lys Glu Lys Asp Leu Ile Arg Thr Ser Glu Ser Asp Cys
65 70 75
Tyr Cys Tyr Asn Gln Asn Ser Gln Val Glu Trp Lys Tyr Ile Trp
80 85 90
Ser Thr Met Gln Val Lys Ile Thr Ser Pro Gly Leu Phe Arg Ile
95 100 105
Val Tyr Ile Ala Glu Arg His Asn Cys Gln Tyr Pro Glu Asn Ile
110 115 120
Leu Ser Phe Ile Lys Cys Val Ile His Asn Phe Trp Ile Pro Lys
125 130 135
Glu Ser Asn Glu Ile Thr Ile Ile Ile Asn Pro Tyr Arg Glu Thr
140 145 150
Val Cys Phe Ser Val Glu Pro Val Lys Lys Ile Phe Asn Tyr Met
155 160 165
Ile His Val Asn Arg Asn Ile Met Asp Phe Lys Leu Phe Leu Val
170 175 180
Phe Val Ala Gly Val Phe Leu Phe Phe Tyr Ala Arg Thr Leu Ser
185 190 195
Gln Ser Pro Thr Phe Tyr Tyr Ser Ser Gly Thr Val Leu Gly Val
200 205 210
Leu Met Thr Leu Val Phe Val Leu Leu Leu Val Lys Arg Phe Ile
215 220 225
Pro Lys Tyr Ser Thr Phe Trp Ala Leu Met Val Gly Cys Trp Phe
230 235 240
Ala Ser Val Tyr Ile Val Cys Gln Leu Met Glu Asp Leu Lys Trp
245 250 255
Leu Trp Tyr Glu Asn Arg Ile Tyr Val Leu Gly Tyr Val Leu Ile
260 265 270
Val Gly Phe Phe Ser Phe Val Val Cys Tyr Lys His Gly Pro Leu
275 280 285
Ala Asp Asp Arg Ser Arg Ser Leu Leu Met Trp Met Leu Arg Leu
290 295 300
Leu Ser Leu Val Leu Val Tyr Ala Gly Val Ala Val Pro Gln Phe
305 310 315
Ala Tyr Ala Ala Ile Ile Leu Leu Met Ser Ser Trp Ser Leu His
320 325 330
Tyr Pro Leu Arg Ala Cys Ser Tyr Met Arg Trp Lys Met Glu Gln
335 340 345
Trp Phe Thr Ser Lys Glu Leu Val Val Lys Tyr Leu Thr Glu Asp
350 355 360
Glu Tyr Arg Glu Gln Ala Asp Ala Glu Thr Asn Ser Ala Leu Glu
365 370 375
Glu Leu Arg Arg Ala Cys Arg Lys Pro Asp Phe Pro Ser Trp Leu
380 385 390
Val Val Ser Arg Leu His Thr Pro Ser Asn

395 400

<210> 30
<211> 133
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4000975CD1

<400> 30
Met Pro Arg Lys Leu Pro Ser Trp Arg Glu Ser Leu Phe Leu Ser
1 5 10 15
Val Glu Leu Ser Pro Leu Ala Leu Ala Met Gly Ser Ala Pro Gly
20 25 30
Leu Gln Val Phe Ser Lys Thr Asn Pro Leu Phe Leu Ser Pro Pro
35 40 45
Leu Lys Ser Arg Ala Leu Gly Pro Ser Pro Gln Glu Gly Phe Trp
50 55 60
Pro Asn Leu Gln Arg Gln Val Arg Ala Val Ser Leu Gly Cys Glu
65 70 75
Ala Ala Gly Glu Gly Asp Phe Gly Gln Met Ser Leu Gly Cys Glu
80 85 90
Ala Ala Gly Glu Gly Asp Phe Gly Gln Met Ser Leu Gly Cys Glu
95 100 105
Ala Ala Gly Glu Gly Asp Phe Gly Gln Val Ser Pro Ala Leu Cys
110 115 120
Pro Ser Gln Val Gln Leu Arg Asp Gly Leu Cys Leu Leu
125 130

<210> 31
<211> 359
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4598831CD1

<400> 31
Met Arg Ala Ala Ala Met Thr Thr Ala Ile Leu Glu Arg Leu Ser
1 5 10 15
Thr Leu Ser Val Ser Gly Gln Gln Leu Arg Arg Leu Pro Lys Ile
20 25 30
Leu Glu Asp Gly Leu Pro Lys Met Pro Cys Thr Val Pro Glu Thr
35 40 45
Asp Val Pro Gln Leu Phe Arg Glu Pro Tyr Ile Arg Thr Gly Tyr
50 55 60
Arg Pro Thr Gly His Glu Trp Arg Tyr Tyr Phe Phe Ser Leu Phe
65 70 75
Gln Lys His Asn Glu Val Val Asn Val Trp Thr His Leu Leu Ala
80 85 90
Ala Leu Ala Val Leu Leu Arg Phe Trp Ala Phe Ala Glu Ala Glu
95 100 105
Ala Leu Pro Trp Ala Ser Thr His Ser Leu Pro Leu Leu Leu Phe
110 115 120
Ile Leu Ser Ser Ile Thr Tyr Leu Thr Cys Ser Leu Leu Ala His
125 130 135
Leu Leu Gln Ser Lys Ser Glu Leu Ser His Tyr Thr Phe Tyr Phe
140 145 150
Val Asp Tyr Val Gly Val Ser Val Tyr Gln Tyr Gly Ser Ala Leu
155 160 165

```
Ala His Phe Phe Tyr Ser Ser Asp Gln Ala Trp Tyr Asp Arg Phe
                170                 175                 180
Trp Leu Phe Phe Leu Pro Ala Ala Ala Phe Cys Gly Trp Leu Ser
                185                 190                 195
Cys Ala Gly Cys Cys Tyr Ala Lys Tyr Arg Tyr Arg Arg Pro Tyr
                200                 205                 210
Pro Val Met Arg Lys Ile Cys Gln Val Val Pro Ala Gly Leu Ala
                215                 220                 225
Phe Ile Leu Asp Ile Ser Pro Val Ala His Arg Val Ala Leu Cys
                230                 235                 240
His Leu Ala Gly Cys Gln Glu Gln Ala Ala Trp Tyr His Thr Leu
                245                 250                 255
Gln Ile Leu Phe Phe Leu Val Ser Ala Tyr Phe Phe Ser Cys Pro
                260                 265                 270
Val Pro Glu Lys Tyr Phe Pro Gly Ser Cys Asp Ile Val Gly His
                275                 280                 285
Gly His Gln Ile Phe His Ala Phe Leu Ser Ile Cys Thr Leu Ser
                290                 295                 300
Gln Leu Glu Ala Ile Leu Leu Asp Tyr Gln Gly Arg Gln Glu Ile
                305                 310                 315
Phe Leu Gln Arg His Gly Pro Leu Ser Val His Met Ala Cys Leu
                320                 325                 330
Ser Phe Phe Phe Leu Ala Ala Cys Ser Ala Ala Thr Ala Ala Leu
                335                 340                 345
Leu Arg His Lys Val Lys Ala Arg Leu Thr Lys Lys Asp Ser
                350                 355
```

<210> 32
<211> 72
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4992201CD1

<400> 32

```
Met Lys Cys Ala Leu Ile Thr Gly Arg Leu Arg Arg Gly Asn Glu
  1               5                  10                  15
Thr Ser Cys Ile Asp His Arg Ala Gln Ser Leu Ala Phe Arg Lys
                20                  25                  30
Pro Ser Val Arg Val His Asp Ala Met Val Ser Val Ile Ile Leu
                35                  40                  45
Phe Ile Leu Ile Ile Thr Phe Ile Ile Phe Leu Leu Phe Leu Glu
                50                  55                  60
Asn Ser Leu Glu Gly Leu Ile Pro Cys Tyr His Gly
                65                  70
```

<210> 33
<211> 112
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5441583CD1

<400> 33

```
Met Asn Ser Ala Met Trp His Gln Thr Ala Thr Gln Thr Thr Val
  1               5                  10                  15
Ser Ser Ala Thr Leu Val Thr His Ile Gln Ala Arg Phe His Leu
                20                  25                  30
Gln Gln Ser Trp Met Arg Trp Leu Ala Glu Ala Asn Pro Leu Pro
```

```
                35                  40                  45
Ala Leu Gln Ala Lys Ala Gly Met Trp Pro Arg Trp Phe Leu Arg
                50                  55                  60
Ser Leu Thr Ile Leu Arg Ser Cys Ile Leu Ser Ile Ser Gly Gln
                65                  70                  75
Arg Cys Leu His Ala Pro Ser Ser Phe Val Ser Leu Met Phe Leu
                80                  85                  90
Ala Thr Cys Tyr Ser Ser Leu Ser Tyr Phe Ser Arg Phe His Arg
                95                 100                 105
Glu Arg Phe Ser Cys Pro Trp
               110
```

<210> 34
<211> 149
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1639243CD1

```
<400> 34
Met Leu Gly Cys Tyr Gly Met Gly Gln Leu Cys Ile Trp Glu Ser
  1               5                  10                  15
Pro Pro Ala Ser Pro Ser Trp Leu Leu Ser Val Gly Cys Tyr His
                20                  25                  30
Leu Pro Ser Leu Gly Leu Leu Ser Pro His Pro Phe Thr Arg Gln
                35                  40                  45
Leu Pro Phe Arg Thr His Trp Pro Ile Pro Ser Phe Ser Ser Ser
                50                  55                  60
His Pro Ser Thr Pro Val His Gly Cys Cys Arg Ser Gly Phe Phe
                65                  70                  75
Val Phe Val Phe Phe Lys Thr Glu Ser His Ser Ala Ala Arg Leu
                80                  85                  90
Glu Cys Ser Gly Arg Ile Leu Ala His Cys Asn Leu Cys Leu Pro
                95                 100                 105
Gly Ser Ser Asp Ser Pro Ala Ser Ala Ser Arg Val Ala Gly Thr
               110                 115                 120
Thr Gly Thr Cys His His Ile Gln Leu Ile Phe Val Phe Leu Val
               125                 130                 135
Glu Met Gly Phe His His Val Gly Gln Asp Leu Leu Thr Ser
               140                 145
```

<210> 35
<211> 97
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1335166CD1

```
<400> 35
Met Glu Ser Cys Ser Val Thr Leu Ala Gly Val Gln Trp Cys Asn
  1               5                  10                  15
Leu Gly Ser Leu Gln Pro Pro Pro Pro Gly Phe Lys Arg Phe Ser
                20                  25                  30
Cys Leu Asn Leu Leu Ser Ser Trp Asp Tyr Arg His Ala Gln Pro
                35                  40                  45
His Trp Leu Phe Phe Val Ser Leu Thr Glu Thr Gly Phe His His
                50                  55                  60
Val Gly Gln Ala Gly Leu Glu Leu Leu Ser Ser Ser Asp Leu Pro
                65                  70                  75
```

```
Ala Leu Ala Ser Gln Ser Ala Gly Ile Thr Gly Val Ser His Cys
                80                  85                  90
Ala Arg Pro Gly Arg Leu Leu
                95

<210> 36
<211> 104
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 166894CD1

<400> 36
Met Phe Ser Glu Ala Leu Leu Ile His Arg Thr Tyr Leu Ala Tyr
  1               5                  10                  15
Leu Phe Ala Cys Leu Leu Leu Met Ser Ser Leu Thr Glu Ser Leu
                20                  25                  30
Leu Gln Arg Thr Thr Pro Ala Ser Arg Pro Arg Asn Val Gly Lys
                35                  40                  45
Gly Lys Ala Trp Leu Val Leu Val Glu Met Glu Met Leu Val Thr
                50                  55                  60
Val Glu Glu Cys Pro Pro Ser Asp Ser Gln Trp Gly Gly Ala Leu
                65                  70                  75
Gly Pro Cys His Cys Pro Arg Thr Ser Ala Phe Gly Cys Pro Ala
                80                  85                  90
Glu Arg Met Arg His Leu Ser Ser Ser Phe Trp Ser Pro Glu
                95                 100

<210> 37
<211> 99
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 217969CD1

<400> 37
Met Ser Ser Leu Cys Val Ser Val Thr Ser Lys Asn His Asn Met
  1               5                  10                  15
Phe Met Ala His Asp Gly Tyr Cys Ser Phe Val Phe Cys Phe Phe
                20                  25                  30
Phe Glu Thr Glu Ser Ala Ser Val Thr Gln Pro Gly Val Gln Trp
                35                  40                  45
Tyr His His Ser Ser Leu Gln Pro Arg Pro Pro Gly Leu Glu Gly
                50                  55                  60
Ser Ser His Leu Ser Leu Gln Val Ala Arg Thr Ile Gly Val Cys
                65                  70                  75
His His Thr Gln Leu Ile Leu Phe Arg Trp Gly Leu Thr Met Leu
                80                  85                  90
Pro Trp Leu Val Ser Asn Phe Arg Ala
                95

<210> 38
<211> 80
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 335237CD1
```

<400> 38
Met Pro Asp Leu Ala Val Val Leu Phe Cys Ser Arg Val Pro Arg
1 5 10 15
Ser Ser Ser Gly Thr Gly Ser Gln Gly Gln Leu Val Pro Arg Ala
20 25 30
Ser Leu Ala Cys Pro Leu Gly Ser Ser Arg Asp Asn Leu Thr Cys
35 40 45
Pro Ile Lys Ala Lys Gly Gln Asn Arg Arg Gln Asn Leu Ala Arg
50 55 60
Pro Ser Ser Asn Ser Lys Gly Lys Pro Val Pro Trp Ile Leu Ser
65 70 75
Glu Ile Lys Thr Lys
80

<210> 39
<211> 96
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 938306CD1

<400> 39
Met Lys Cys Lys Gly Ile Leu Ser Val Pro Gly Trp Leu Pro Thr
1 5 10 15
Val Leu Gly Lys Arg Val Ile Phe Gln Lys Gly Pro Glu Gln Ser
20 25 30
Ala Cys Ile Leu Ser Pro Leu Leu Pro Val Ser Ser Lys Ala Ser
35 40 45
Gln Lys Leu His Phe Pro Thr Ser Cys His Phe Gln Asn His Ser
50 55 60
Leu Asn Leu Lys Asn Lys Trp Glu Ala Val Phe Leu Pro Leu Met
65 70 75
Ile Ala Ala Thr Tyr Lys Pro Ala Arg Thr Glu His Ser Lys Gln
80 85 90
Arg Arg Val Gln Ser Cys
95

<210> 40
<211> 92
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1448129CD1

<400> 40
Met Ser Ala Met Phe Asn Ala Pro Trp Trp Ser Leu Gly Lys Met
1 5 10 15
Pro Thr Pro Tyr Leu Leu Ser Leu Met Asn Ser Gln Ala Ser Phe
20 25 30
Gly Gln Thr Phe Gln Gln Ala Leu Glu Ser Arg Leu Ile Val Thr
35 40 45
Arg Glu Arg Tyr Lys Leu Gly Glu Arg Lys Glu Pro Phe Leu Glu
50 55 60
Glu Ser Ala Phe Glu Gln Phe Leu Lys Val Leu Val Gly Arg Gly
65 70 75
His Ser Arg Gln Val Gly Leu Phe Thr Glu Trp Thr Ala Val Trp
80 85 90
Val Ala

<210> 41
<211> 77
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1761049CD1

<400> 41
Met Ile Asn Val Trp Tyr His Val Phe Leu Gln Asn Ile Glu Phe
1 5 10 15
Lys Glu Cys Ser Leu Gln Tyr Trp Gln Leu Ser Pro Asp Leu Leu
20 25 30
Phe Asn His Gly Val Ile Ser Glu Lys Tyr Leu Phe Tyr Phe Ile
35 40 45
Leu Phe Tyr Phe Ile Leu Phe Met Leu Phe Met Leu Phe Met Leu
50 55 60
Cys Tyr Val Met Leu Cys Tyr Val Met Leu Cys Tyr Val Met Leu
65 70 75
Phe Phe

<210> 42
<211> 75
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1959587CD1

<400> 42
Met Lys Leu Pro His Leu Ala Gln Phe Leu Thr Ser Pro Leu Val
1 5 10 15
Leu Trp Ser Thr Gly Val Ser Gly Ser Ala Gly Phe His Gln Leu
20 25 30
Val Pro Gln Trp Glu Cys Glu Glu Val Pro Gly Cys Gly Lys Ser
35 40 45
Cys Leu Ser Lys Arg Gly Leu Ile Glu Met Leu Gly Lys Val Ala
50 55 60
Val Ser Leu His Tyr Gly Arg Glu Gln Ser Gly Arg Ala Cys Cys
65 70 75

<210> 43
<211> 85
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2303463CD1

<400> 43
Met Ala Ile Phe Ser Leu Leu Met Phe His Ile Tyr Ser Phe Met
1 5 10 15
Arg Ile Phe Ser Phe Ala Leu Met Ser Val Phe Ile Ile Ala Ala
20 25 30
Phe Lys Phe Leu Ser Ala Val Tyr Ile Leu Asp Ile Leu Glu Met
35 40 45
Ala Thr Ala Cys Phe Leu Ser Cys Val Phe Ile Thr Phe Ser Arg
50 55 60

```
Val Phe Thr His Leu Leu Asn Trp Lys Leu Cys Pro Gly Asp Cys
                65                  70                  75
Ile Gln Asp Trp Ile Lys Lys Thr Gly Phe
                80                  85

<210> 44
<211> 89
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2512281CD1

<400> 44
Met Ala Ala Ala Pro Ala Pro Lys Pro Ser Leu Ala Pro Val Leu
  1                 5                  10                  15
Gly Pro Leu Glu Val Leu Pro Ala Pro Leu Gln Ala Pro Thr Arg
                20                  25                  30
Arg Ser Pro Gly Thr Glu Cys Ala Pro Pro Ala Thr Gly Lys Gly
                35                  40                  45
Arg Leu Ile Arg Val Arg Ser Arg Asp Gly Ile Val Thr Met Lys
                50                  55                  60
Ser Ser Arg Arg Ala Met Cys Leu Lys Pro Ser Val Thr Leu Pro
                65                  70                  75
Asn Ser Gln Glu Ala Arg His Ala Leu His Pro Ala Glu Pro
                80                  85

<210> 45
<211> 123
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2755924CD1

<400> 45
Met Phe Ile Lys Ile His Asn Leu Phe Phe Cys Ile Cys Val Leu
  1                 5                  10                  15
Pro Thr Leu Ala Ile Ser Gly Trp Ser Cys Pro Ser Leu Leu Ser
                20                  25                  30
Leu Ser Phe Phe Lys His Ser Ile Cys Ile Leu Phe Leu Phe Leu
                35                  40                  45
Val Thr Gly Phe His Tyr Val Ala His Thr Gly His Glu Leu Leu
                50                  55                  60
Ser Ser Gly Asp Leu Pro Thr Ser Ala Ser Gln Val Ala Gly Thr
                65                  70                  75
Thr Gly Thr Cys His Cys Ala Gln Leu Val Thr Ala Asn Phe Asn
                80                  85                  90
Leu Gly Met Phe Val Pro Leu Leu Tyr Cys His Val Lys Asn Phe
                95                 100                 105
Ala Asn Ser Gln Glu Thr Ser Val Ser Ser Val Lys Leu Asn Leu
               110                 115                 120
Ser Ser Leu


<210> 46
<211> 159
<212> PRT
<213> Homo sapiens

<220>
```

<221> misc_feature
<223> Incyte ID No: 2796369CD1

<400> 46
Met Gly Ser Leu Cys Ser Glu Asn Gly Arg Val Trp Asp Gly Leu
1 5 10 15
Ser Phe Leu Leu Val Gly Pro Gly Ser Gly Ser Gly Ala Ala Pro
20 25 30
Phe Leu Trp Ser Thr Gln Arg Glu Gln Glu Gly Leu Asp Leu Gly
35 40 45
Lys Glu Ala Ile His Arg Ala Pro Gln Lys Pro Gly Pro Pro Gly
50 55 60
Ala His Cys Cys Ala Glu Ala Thr Arg Leu Gly Tyr Phe Leu Pro
65 70 75
Glu Ala Gly Asn Arg Glu Cys Arg Glu Ala Arg Gln Gln Gln Glu
80 85 90
Ala Pro Asn Ala Gly Val Ser Lys Pro Glu Pro Pro Pro Asp Phe
95 100 105
Thr Pro Val Cys Pro Ala His Ser Arg Leu Ser Leu Gly Gly Pro
110 115 120
Trp Gly Leu Asp Leu Pro Asp Leu Trp Pro Gln Lys Gly Leu Ser
125 130 135
Pro Glu Ser His Gly Met Glu Pro Gly Met His Arg Pro Ser Gly
140 145 150
Leu Cys Leu Gly Ser Arg Pro Gly Ile
155

<210> 47
<211> 77
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3010920CD1

<400> 47
Met Ala Leu Arg Lys Ser Ser Cys Leu Pro Leu Lys Leu Gly Thr
1 5 10 15
Leu Ile Thr Tyr Ser Leu Ile Phe Leu Ala Trp Phe Leu Leu Lys
20 25 30
Ser Ala Thr Phe Asn Gln Val Ile Met Pro Arg Glu Leu Cys Gln
35 40 45
Asp Leu Ile Tyr Val His Ser Tyr Asp Lys Tyr Leu Leu Ile Phe
50 55 60
Gln Ile Asn Ser Cys Gly Cys Cys Asn Thr Tyr Ile His Tyr Arg
65 70 75
Lys Leu

<210> 48
<211> 130
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3360955CD1

<400> 48
Met Leu Phe Val Phe Ser Phe Cys Pro Gln Gln Ala Val Thr Ser
1 5 10 15
Asp Gln Glu Val Ser Lys Ser Thr Glu Thr Leu Arg Arg Leu Met

```
                20                  25                  30
Leu Ser Ala Lys Ile Met Asp Gly Glu Asp Thr Gly Leu Tyr His
                35                  40                  45
Gln His Phe Ser Trp Tyr Leu Thr Ile Asn Arg Met Met Ala His
                50                  55                  60
Arg Ser Lys Gly Thr Ser Phe His Ala Leu Pro Ser Leu Pro Ile
                65                  70                  75
Leu Ala Asn Pro Ser Ser Trp Pro Pro Asp Tyr Asp Thr Thr Gln
                80                  85                  90
Met Ser Ile Phe Ser Ala Arg Lys Ser Leu Leu Gly Thr Lys Leu
                95                 100                 105
Leu Thr Ser Cys Leu Ser Ser Leu His Phe Arg Lys Cys Pro Val
               110                 115                 120
Leu His Cys Asn Leu Leu Lys Ala Gly Lys
               125                 130
```

<210> 49
<211> 97
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3409459CD1

<400> 49

```
Met Asn Phe Tyr Arg Ala Ser Cys Leu Ser Leu Trp Val Phe Ala
  1                 5                  10                  15
Gly Gly Gly Phe Gly Leu Asn Ala Ala Asp Met Ser Asp Ser Pro
                20                  25                  30
Leu Ala Ala Ala Gly Glu Val Ala Ile Val Val Pro Leu His Pro
                35                  40                  45
Gly His Leu Arg Cys Trp Tyr Leu Leu Asn Gln Gly Ile Trp Pro
                50                  55                  60
Gly Arg Ala Ser Ser Pro Ala Pro Pro Ala Trp His Cys Pro Leu
                65                  70                  75
Pro Val Leu Gln Arg Ala Ile Arg Lys Ala Gly Leu Pro Thr Leu
                80                  85                  90
Leu Pro Arg Pro Ala Gly Pro
                95
```

<210> 50
<211> 74
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4102938CD1

<400> 50

```
Met Pro Ser Leu Leu Asp His Pro Phe Ala Glu Lys Pro Phe Leu
  1                 5                  10                  15
Leu Leu Ala Leu Phe Gln Leu Asn Phe Leu Ala Pro Leu Ser Gln
                20                  25                  30
Val Ala Gly His Ala Ala Glu Gly Asn Trp Gly Asp Ser Arg Thr
                35                  40                  45
Ala Asn His Phe Ser Lys Leu Arg Phe Gln Phe Glu Thr Arg Leu
                50                  55                  60
Ala Asn Met Val Lys Pro Arg Leu Tyr Lys Lys Tyr Lys Asn
                65                  70
```

<210> 51

<211> 74
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4124601CD1

<400> 51
Met Leu Gly Leu Gln Gln Gly Gln Gly Gly Ser Ser Glu Arg Gln
1 5 10 15
Lys Trp Val Gly Pro Arg Gly Trp Arg Ala Ala Glu His Lys Ser
20 25 30
Arg Leu Lys Gly Ala Ala Thr Ala Gln Ser Pro Leu Thr Ala Ala
35 40 45
Gly Trp Asp Cys Lys Pro Arg Val Ala Arg Ser Val Ser Phe Phe
50 55 60
Gln Asp Lys Leu Glu Ile Arg Phe Ser His Gly Ile Val Ser
65 70

<210> 52
<211> 151
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4180577CD1

<400> 52
Met Ser Ser Ser Thr Ser Phe Ile Leu Ser Ala Ile Ala Ser Gly
1 5 10 15
Phe His Tyr Ser Leu Ser Ala Val Thr Ala Cys Gly Gln Leu Leu
20 25 30
Leu Leu Thr Ala Cys Arg Glu Leu Pro Asn Phe Ser Ser Gln Phe
35 40 45
Phe Leu Arg Ser Trp Leu Phe Trp Pro Gln Leu Lys Gly Val Leu
50 55 60
Leu Ser Ser Leu Arg Val Leu Ser Leu Phe Asp Pro Ile Val Val
65 70 75
Phe Ser Ser Phe Glu His Val Phe Gln Tyr Ser Tyr Phe Asn Leu
80 85 90
Leu Arg Thr Leu Lys Gly Asn Asp Lys Leu Val Val Gly Ile Trp
95 100 105
Gln Thr Gly Ala Cys Leu Phe Glu Arg Ser Ser Arg Arg Asp Lys
110 115 120
Ile Gln Ser Ala Ile Cys Phe Ser Trp Arg Gly Lys Arg Glu Asn
125 130 135
Leu Leu Asp Tyr Ile Leu Val Pro Trp His Thr Thr Tyr Met Phe
140 145 150
Lys

<210> 53
<211> 137
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5265807CD1

<400> 53

```
Met Leu Thr Tyr Ser Ser Phe His Phe Leu Leu Phe Tyr Leu Leu
 1               5                  10                  15
Leu Pro Leu Ser Leu Leu Ser Pro Ala Pro Gln Gln Lys Val Leu
            20                  25                  30
Gly Leu Leu Leu Ala His Ser Ala Asp Val Asn Ala Arg Asp Lys
            35                  40                  45
Leu Trp Gln Thr Pro Leu His Val Ala Ala Ala Asn Arg Ala Thr
            50                  55                  60
Lys Cys Ala Glu Ala Leu Ala Pro Leu Leu Ser Ser Leu Asn Val
            65                  70                  75
Ala Asp Arg Ser Gly Arg Ser Ala Leu His His Ala Val His Ser
            80                  85                  90
Gly His Leu Glu Val Arg Thr Val Pro Ile Gln Ala Gln Leu Gly
            95                  100                 105
Leu Ser Leu Phe Leu Pro Ser Tyr Ser Arg Phe Pro Ala Ser Gly
            110                 115                 120
Pro Ser Ser Leu Lys Glu Lys Gln Pro Gly Trp Leu Tyr Lys His
            125                 130                 135
Leu Ser
```

<210> 54
<211> 137
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5405979CD1

<400> 54

```
Met Gln Ser Phe Thr Phe Tyr Leu Val Leu Pro Ser Pro Val Val
 1               5                  10                  15
Leu Ala Pro Pro Val Pro Ser Ala Ala Gly Pro Val Phe Ser Phe
            20                  25                  30
Gln Pro Arg Ser Ser Gln Pro Leu Leu His Gln Trp Cys Leu Leu
            35                  40                  45
Trp Ala Ser Pro Arg Leu Arg Cys Phe Arg Leu Ser Leu Leu Arg
            50                  55                  60
Gln Gln His Ala Ser Arg Trp His Ala Cys Pro Leu His Ala Ser
            65                  70                  75
Leu Gly Leu Pro Leu Leu Ala Gly Gln Gln Pro Ala Glu Pro Arg
            80                  85                  90
Tyr Leu Pro Phe Pro Cys Cys Ser Ser Leu Ser Pro Leu Ser Ser
            95                  100                 105
Trp Ala Cys Leu Gly Gln Lys Gly Gln Val Ser Gly Thr Ser Gln
            110                 115                 120
Glu Thr Leu Gly Arg Glu Val Ser Leu Ser Leu Glu Thr Val Asp
            125                 130                 135
Lys Leu
```

<210> 55
<211> 205
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7481109CD1

<400> 55

```
Met Met Arg Thr Leu Ile Thr Thr His Pro Leu Pro Leu Leu Leu
```

```
 1               5                  10                  15
Leu Pro Gln Gln Leu Leu Gln Leu Val Gln Phe Gln Glu Val Asp
            20                  25                  30
Thr Asp Phe Asp Phe Pro Glu Glu Asp Lys Lys Glu Glu Phe Glu
            35                  40                  45
Glu Cys Leu Glu Lys Phe Phe Ser Thr Gly Pro Ala Arg Pro Pro
            50                  55                  60
Thr Lys Glu Lys Val Lys Arg Arg Val Leu Ile Glu Pro Gly Met
            65                  70                  75
Pro Leu Asn His Ile Glu Tyr Cys Asn His Glu Ile Met Gly Lys
            80                  85                  90
Asn Val Tyr Tyr Lys His Arg Trp Val Ala Glu His Tyr Phe Leu
            95                  100                 105
Leu Met Gln Tyr Asp Glu Leu Gln Lys Ile Cys Tyr Asn Arg Phe
            110                 115                 120
Val Pro Cys Lys Asn Gly Ile Arg Lys Cys Asn Arg Ser Lys Gly
            125                 130                 135
Leu Val Glu Gly Val Tyr Cys Asn Leu Thr Glu Ala Phe Glu Ile
            140                 145                 150
Pro Ala Cys Lys Tyr Glu Ser Leu Tyr Arg Lys Gly Tyr Val Leu
            155                 160                 165
Ile Thr Cys Ser Trp Gln Asn Glu Met Gln Lys Arg Ile Pro His
            170                 175                 180
Thr Ile Asn Asp Leu Val Glu Pro Pro Glu His Arg Ser Phe Leu
            185                 190                 195
Ser Glu Asp Gly Val Phe Val Ile Ser Pro
            200                 205
```

<210> 56
<211> 199
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6247114CD1

<400> 56

```
Met Glu Thr Phe Pro Leu Leu Leu Leu Ser Leu Gly Leu Val Leu
 1               5                  10                  15
Ala Glu Ala Ser Glu Ser Thr Met Lys Ile Ile Lys Glu Glu Phe
            20                  25                  30
Thr Asp Glu Glu Met Gln Tyr Asp Met Ala Lys Ser Gly Gln Glu
            35                  40                  45
Lys Gln Thr Ile Glu Ile Leu Met Asn Pro Ile Leu Leu Val Lys
            50                  55                  60
Asn Thr Ser Leu Ser Met Ser Lys Asp Asp Met Ser Ser Thr Leu
            65                  70                  75
Leu Thr Phe Arg Ser Leu His Tyr Asn Asp Pro Lys Gly Asn Ser
            80                  85                  90
Ser Gly Asn Asp Lys Glu Cys Cys Asn Asp Met Thr Val Trp Arg
            95                  100                 105
Lys Val Ser Glu Ala Asn Gly Ser Cys Lys Trp Ser Asn Asn Phe
            110                 115                 120
Ile Arg Ser Ser Thr Glu Val Met Arg Arg Val His Arg Ala Pro
            125                 130                 135
Ser Cys Lys Phe Val Gln Asn Pro Gly Ile Ser Cys Cys Glu Ser
            140                 145                 150
Leu Glu Leu Glu Asn Thr Val Cys Gln Phe Thr Thr Gly Lys Gln
            155                 160                 165
Phe Pro Arg Cys Gln Tyr His Ser Val Thr Ser Leu Glu Lys Ile
            170                 175                 180
Leu Thr Val Leu Thr Gly His Ser Leu Met Ser Trp Leu Val Cys
```

185 190 195
Gly Ser Lys Leu

<210> 57
<211> 719
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3243866CD1

<400> 57
Met Glu Lys Ile Leu Phe Tyr Leu Phe Leu Ile Gly Ile Ala Val
1 5 10 15
Lys Ala Gln Ile Cys Pro Lys Arg Cys Val Cys Gln Ile Leu Ser
20 25 30
Pro Asn Leu Ala Thr Leu Cys Ala Lys Lys Gly Leu Leu Phe Val
35 40 45
Pro Pro Asn Ile Asp Arg Arg Thr Val Glu Leu Arg Leu Ala Asp
50 55 60
Asn Phe Val Thr Asn Ile Lys Arg Lys Asp Phe Ala Asn Met Thr
65 70 75
Ser Leu Val Asp Leu Thr Leu Ser Arg Asn Thr Ile Ser Phe Ile
80 85 90
Thr Pro His Ala Phe Ala Asp Leu Arg Asn Leu Arg Ala Leu His
95 100 105
Leu Asn Ser Asn Arg Leu Thr Lys Ile Thr Asn Asp Met Phe Ser
110 115 120
Gly Leu Ser Asn Leu His His Leu Ile Leu Asn Asn Asn Gln Leu
125 130 135
Thr Leu Ile Ser Ser Thr Ala Phe Asp Asp Val Phe Ala Leu Glu
140 145 150
Glu Leu Asp Leu Ser Tyr Asn Asn Leu Glu Thr Ile Pro Trp Asp
155 160 165
Ala Val Glu Lys Met Val Ser Leu His Thr Leu Ser Leu Asp His
170 175 180
Asn Met Ile Asp Asn Ile Pro Lys Gly Thr Phe Ser His Leu His
185 190 195
Lys Met Thr Arg Leu Asp Val Thr Ser Asn Lys Leu Gln Lys Leu
200 205 210
Pro Pro Asp Pro Leu Phe Gln Arg Ala Gln Val Leu Ala Thr Ser
215 220 225
Gly Ile Ile Ser Pro Ser Thr Phe Ala Leu Ser Phe Gly Gly Asn
230 235 240
Pro Leu His Cys Asn Cys Glu Leu Leu Trp Leu Arg Arg Leu Ser
245 250 255
Arg Glu Asp Asp Leu Glu Thr Cys Ala Ser Pro Pro Leu Leu Thr
260 265 270
Gly Arg Tyr Phe Trp Ser Ile Pro Glu Glu Glu Phe Leu Cys Glu
275 280 285
Pro Pro Leu Ile Thr Arg His Thr His Glu Met Arg Val Leu Glu
290 295 300
Gly Gln Arg Ala Thr Leu Arg Cys Lys Ala Arg Gly Asp Pro Glu
305 310 315
Pro Ala Ile His Trp Ile Ser Pro Glu Gly Lys Leu Ile Ser Asn
320 325 330
Ala Thr Arg Ser Leu Val Tyr Asp Asn Gly Thr Leu Asp Ile Leu
335 340 345
Ile Thr Thr Val Lys Asp Thr Gly Ala Phe Thr Cys Ile Ala Ser
350 355 360
Asn Pro Ala Gly Glu Ala Thr Gln Ile Val Asp Leu His Ile Ile

```
                365                 370                 375
Lys Leu Pro His Leu Leu Asn Ser Thr Asn His Ile His Glu Pro
                380                 385                 390
Asp Pro Gly Ser Ser Asp Ile Ser Thr Ser Thr Lys Ser Gly Ser
                395                 400                 405
Asn Thr Ser Ser Ser Asn Gly Asp Thr Lys Leu Ser Gln Asp Lys
                410                 415                 420
Ile Val Val Ala Glu Ala Thr Ser Ser Thr Ala Leu Leu Lys Phe
                425                 430                 435
Asn Phe Gln Arg Asn Ile Pro Gly Ile Arg Met Phe Gln Ile Gln
                440                 445                 450
Tyr Asn Gly Thr Tyr Asp Asp Thr Leu Val Tyr Arg Met Ile Pro
                455                 460                 465
Pro Thr Ser Lys Thr Phe Leu Val Asn Asn Leu Ala Ala Gly Thr
                470                 475                 480
Met Tyr Asp Leu Cys Val Leu Ala Ile Tyr Asp Asp Gly Ile Thr
                485                 490                 495
Ser Leu Thr Ala Thr Arg Val Val Gly Cys Ile Gln Phe Thr Thr
                500                 505                 510
Glu Gln Asp Tyr Val Arg Cys His Phe Met Gln Ser Gln Phe Leu
                515                 520                 525
Gly Gly Thr Met Ile Ile Ile Ile Gly Gly Ile Ile Val Ala Ser
                530                 535                 540
Val Leu Val Phe Ile Ile Ile Leu Met Ile Arg Tyr Lys Val Cys
                545                 550                 555
Asn Asn Asn Gly Gln His Lys Val Thr Lys Val Ser Asn Val Tyr
                560                 565                 570
Ser Gln Thr Asn Gly Ala Gln Ile Gln Gly Cys Ser Val Thr Leu
                575                 580                 585
Pro Gln Ser Val Ser Lys Gln Ala Val Gly His Glu Glu Asn Ala
                590                 595                 600
Gln Cys Cys Lys Ala Thr Ser Asp Asn Val Ile Gln Ser Ser Glu
                605                 610                 615
Thr Cys Ser Ser Gln Asp Ser Ser Thr Thr Thr Ser Ala Leu Pro
                620                 625                 630
Pro Ser Trp Thr Ser Ser Thr Ser Val Ser Gln Lys Gln Lys Arg
                635                 640                 645
Lys Thr Gly Thr Lys Pro Ser Thr Glu Pro Gln Asn Glu Ala Val
                650                 655                 660
Thr Asn Val Glu Ser Gln Asn Thr Asn Arg Asn Asn Ser Thr Ala
                665                 670                 675
Leu Gln Leu Ala Ser Arg Pro Pro Asp Ser Val Thr Glu Gly Pro
                680                 685                 690
Thr Ser Lys Arg Ala His Ile Lys Pro Asn Ala Leu Leu Thr Asn
                695                 700                 705
Val Asp Gln Ile Val Gln Glu Thr Gln Arg Leu Glu Leu Ile
                710                 715
```

<210> 58
<211> 383
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7475633CD1

<400> 58

```
Met Pro Ser Gly Cys Arg Cys Leu His Leu Val Cys Leu Leu Cys
  1               5                  10                  15
Ile Leu Gly Ala Pro Gly Gln Pro Val Arg Ala Asp Asp Cys Ser
             20                  25                  30
Ser His Cys Asp Leu Ala His Gly Cys Cys Ala Pro Asp Gly Ser
```

```
                35                  40                  45
Cys Arg Cys Asp Pro Gly Trp Glu Gly Leu His Cys Glu Arg Cys
                50                  55                  60
Val Arg Met Pro Gly Cys Gln His Gly Thr Cys His Gln Pro Trp
                65                  70                  75
Gln Cys Ile Cys His Ser Gly Trp Ala Gly Lys Phe Cys Asp Lys
                80                  85                  90
Asp Glu His Ile Cys Thr Thr Gln Ser Pro Cys Gln Asn Gly Gly
                95                  100                 105
Gln Cys Met Tyr Asp Gly Gly Gly Glu Tyr His Cys Val Cys Leu
               110                 115                 120
Pro Gly Phe His Gly Arg Asp Cys Glu Arg Lys Ala Gly Pro Cys
               125                 130                 135
Glu Gln Ala Gly Ser Pro Cys Arg Asn Gly Gly Gln Cys Gln Asp
               140                 145                 150
Asp Gln Gly Phe Ala Leu Asn Phe Thr Cys Arg Cys Leu Val Gly
               155                 160                 165
Phe Val Gly Ala Arg Cys Glu Val Asn Val Asp Asp Cys Leu Met
               170                 175                 180
Arg Pro Cys Ala Asn Gly Ala Thr Cys Leu Asp Gly Ile Asn Arg
               185                 190                 195
Phe Ser Cys Leu Cys Pro Glu Gly Phe Ala Gly Arg Phe Cys Thr
               200                 205                 210
Ile Asn Leu Asp Asp Cys Ala Ser Arg Pro Cys Gln Arg Gly Ala
               215                 220                 225
Arg Cys Arg Asp Arg Val His Asp Phe Asp Cys Leu Cys Pro Ser
               230                 235                 240
Gly Tyr Gly Gly Lys Thr Cys Glu Leu Val Leu Pro Val Pro Asp
               245                 250                 255
Pro Pro Thr Thr Val Asp Thr Pro Leu Gly Pro Thr Ser Ala Val
               260                 265                 270
Val Val Pro Ala Thr Gly Pro Ala Pro His Ser Ala Gly Ala Gly
               275                 280                 285
Leu Leu Arg Ile Ser Val Lys Glu Val Val Arg Arg Gln Glu Ala
               290                 295                 300
Gly Leu Gly Glu Pro Ser Leu Val Ala Leu Val Val Phe Gly Ala
               305                 310                 315
Leu Thr Ala Ala Leu Val Leu Ala Thr Val Leu Leu Thr Leu Arg
               320                 325                 330
Ala Trp Arg Arg Gly Val Cys Pro Pro Gly Pro Cys Cys Tyr Pro
               335                 340                 345
Ala Pro His Tyr Ala Pro Ala Cys Gln Asp Gln Glu Cys Gln Val
               350                 355                 360
Ser Met Leu Pro Ala Gly Leu Pro Leu Pro Arg Asp Leu Pro Pro
               365                 370                 375
Glu Pro Gly Lys Thr Thr Ala Leu
               380
```

<210> 59
<211> 126
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1431268CD1

<400> 59

```
Met Ser Arg Leu Glu Ser Ser Glu Ala Ala Cys Arg Ala Val Pro
  1               5                  10                  15
Ser Ala Trp His Thr Phe Leu Leu Ser Pro Leu Cys Leu Leu Leu
                20                  25                  30
Ile Gln Val Trp Ala His Gly Pro Ser Leu Gln Val Val Thr Lys
```

```
                35                  40                  45
Val Ala Pro Pro Ala Leu Thr Ser Ser Met Ser Asp Ser Leu Val
                50                  55                  60
Phe Thr Lys His Phe Ser Leu Cys Lys Val Ile Asp Ser Ala Asn
                65                  70                  75
Val His Arg Gly Cys Thr Thr Cys Gln Ala Leu Val Lys Ala Arg
                80                  85                  90
Asp Val Glu Thr Ser Cys Cys Arg Phe Ser Ala His Ala Leu Ala
                95                  100                 105
Gly Glu Ala Val Ser Gln Gln Asn Lys Gln Arg Gly Gly Glu Ala
                110                 115                 120
Ala Ser Cys Leu Leu Arg
                125
```

<210> 60
<211> 137
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2414185CD1

<400> 60

```
Met Met Gly Lys Leu Ser Pro Thr Phe Ile Leu Gly Ile Cys Trp
  1               5                  10                  15
Val Pro Ala Gly Leu Gly Tyr Gln Gln Gly Lys Lys Thr Trp Pro
                20                  25                  30
Leu Ser Ser Ser Pro Tyr Asn Leu Gln Asp Lys Met Tyr Ala Leu
                35                  40                  45
Glu Lys Ala Gly Asp Pro Ser Lys Ala Arg Ser Met Gly Pro His
                50                  55                  60
Lys Ser Pro Glu Thr Gln Arg Gly Gln Pro Met Glu Met Ser Gly
                65                  70                  75
Leu Lys Gly Gln Val Thr Ser Thr Ala Leu His Thr Leu His Phe
                80                  85                  90
Pro Arg Arg Pro Pro Ser Gly Cys Gln Thr Asp Gln Ala Gly Asp
                95                  100                 105
His Glu Pro Gly Gly Arg Phe Leu Ala Gln Pro Gln Arg Leu Arg
                110                 115                 120
Glu Leu Ser Leu Met Ile Ser Pro Leu Gln Leu Leu Pro Phe Gly
                125                 130                 135
Ser Arg
```

<210> 61
<211> 77
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5266594CD1

<400> 61

```
Met Leu Glu Gly Thr Leu Pro Leu Pro Thr Val Leu Leu Val Gly
  1               5                  10                  15
Arg Pro Val Leu Leu Leu Ala Leu Gly Ala Ala Val Pro Gly His
                20                  25                  30
Leu Ala Ala Pro Thr Asp Val Glu Leu Pro Glu Leu Leu Leu Asn
                35                  40                  45
His Cys Ala Gly Arg Val Val Ala Leu Ile Val Gly Ala Arg Val
                50                  55                  60
```

Leu Leu Leu Phe His Tyr Ile Pro Val Gly Ile Ile Ile Pro Gly
65 70 75
His Ile

<210> 62
<211> 110
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7610617CD1

<400> 62
Met Ser Asn Leu Trp Leu Leu Val Gly Ala Arg Ala Cys Ser Leu
1 5 10 15
Ser Leu Leu Thr Tyr Ser Phe Leu Gly Asp Leu Ile Pro Ser His
20 25 30
Cys Leu Lys His Leu Pro Gly Thr Gly Val Ile His Leu Cys Ser
35 40 45
Ser Ser Ser Glu Ile Pro Ser Ala Pro Phe Ile His Leu Phe Ile
50 55 60
His Ser Ala Asn Ile Cys Gly Ile Ser Val Pro Gly Thr Ala Leu
65 70 75
Gln Pro Gly Cys Thr Ile Gly Thr Gln Thr Asp Thr Pro Phe Pro
80 85 90
Met His Ser Leu Leu Thr Asp Thr Pro Ala Trp Gln Cys Leu Gly
95 100 105
Val Phe Thr Ala Pro
110

<210> 63
<211> 103
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1902436CD1

<400> 63
Met Thr Phe Ile Tyr Thr Phe Ile Leu Ser Phe Phe Leu Gln Leu
1 5 10 15
Cys Cys Ser Phe Met Lys Leu Ile Leu Leu Ile Ser Asn Thr Asn
20 25 30
Ala Val Ser Phe Ile Leu His Arg Pro Cys Thr Leu Cys Ser Asp
35 40 45
Phe Tyr Ser His Ile Cys Met Leu Leu Thr Val Ser Val Asn Phe
50 55 60
Leu Ser Phe Trp Asn Asn Phe Gln Thr Ile Leu Thr Trp Ala Asp
65 70 75
Leu Phe Ser Met Leu Leu Ala Tyr Glu Tyr Arg Phe Thr Arg Leu
80 85 90
Phe Ser Val Leu Pro His Thr Ser Val Met Leu Cys Phe
95 100

<210> 64
<211> 192
<212> PRT
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 2310369CD1

<400> 64
Met Ala Pro Lys Pro Gly Ala Glu Trp Ser Thr Ala Leu Ser His
1 5 10 15
Leu Val Leu Gly Val Val Ser Leu His Ala Ala Val Ser Thr Ala
20 25 30
Glu Ala Ser Arg Gly Ala Ala Ala Gly Phe Leu Leu Gln Val Leu
35 40 45
Ala Ala Thr Thr Thr Leu Ala Pro Gly Leu Ser Thr His Glu Asp
50 55 60
Cys Leu Ala Gly Ala Trp Val Ala Thr Val Ile Gly Leu Pro Leu
65 70 75
Leu Ala Phe Asp Phe His Trp Val Asn Gly Asp Arg Ser Ser Ala
80 85 90
Asn Leu Leu Leu Gly Gly Gly Met Val Leu Ala Val Ala Gly Gly
95 100 105
His Leu Gly Pro Glu Gly Arg Ser Val Ala Gly Gln Ala Met Leu
110 115 120
Leu Val Val Ala Val Thr Ile Leu Ile Val Ala Val Phe Thr Ala
125 130 135
Asn Thr Tyr Gly Met Trp Gly Gly Ala Met Leu Gly Val Ala Gly
140 145 150
Leu Leu Ser Arg Leu Glu Glu Asp Arg Leu Leu Leu Leu Pro Lys
155 160 165
Glu Asp Val Cys Arg Trp Ala Leu Ala Val Gly Ser Trp Ala Tyr
170 175 180
Cys Arg Ala Leu His Thr Gln Arg Leu Gln Trp Glu
185 190

<210> 65
<211> 310
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6180576CD1

<400> 65
Met Asn Gly Leu Ser Leu Ser Glu Leu Cys Cys Leu Phe Cys Cys
1 5 10 15
Pro Pro Cys Pro Gly Arg Ile Ala Ala Lys Leu Ala Phe Leu Pro
20 25 30
Pro Glu Ala Thr Tyr Ser Leu Val Pro Glu Pro Glu Pro Gly Pro
35 40 45
Gly Gly Ala Gly Ala Ala Pro Leu Gly Thr Leu Arg Ala Ser Ser
50 55 60
Gly Ala Pro Gly Arg Trp Lys Leu His Leu Thr Glu Arg Ala Asp
65 70 75
Phe Gln Tyr Ser Gln Arg Glu Leu Asp Thr Ile Glu Val Phe Pro
80 85 90
Thr Lys Ser Ala Arg Gly Asn Arg Val Ser Cys Met Tyr Val Arg
95 100 105
Cys Val Pro Gly Ala Arg Tyr Thr Val Leu Phe Ser His Gly Asn
110 115 120
Ala Val Asp Leu Gly Gln Met Ser Ser Phe Tyr Ile Gly Leu Gly
125 130 135
Ser Arg Leu His Cys Asn Ile Phe Ser Tyr Asp Tyr Ser Gly Tyr
140 145 150
Gly Ala Ser Ser Gly Arg Pro Ser Glu Arg Asn Leu Tyr Ala Asp
155 160 165

```
Ile Asp Ala Ala Trp Gln Ala Leu Arg Thr Arg Tyr Gly Ile Ser
                170                 175                 180
Pro Asp Ser Ile Ile Leu Tyr Gly Gln Ser Ile Gly Thr Val Pro
                185                 190                 195
Thr Val Asp Leu Ala Ser Arg Tyr Glu Cys Ala Ala Val Val Leu
                200                 205                 210
His Ser Pro Leu Thr Ser Gly Met Arg Val Ala Phe Pro Asp Thr
                215                 220                 225
Lys Lys Thr Tyr Cys Phe Asp Ala Phe Pro Asn Ile Glu Lys Val
                230                 235                 240
Ser Lys Ile Thr Ser Pro Val Leu Ile Ile His Gly Thr Glu Asp
                245                 250                 255
Glu Val Ile Asp Phe Ser His Gly Leu Ala Leu Tyr Glu Arg Cys
                260                 265                 270
Pro Lys Ala Val Glu Pro Leu Trp Val Glu Gly Ala Gly His Asn
                275                 280                 285
Asp Ile Glu Leu Tyr Ser Gln Tyr Leu Glu Arg Leu Arg Arg Phe
                290                 295                 300
Ile Ser Gln Glu Leu Pro Ser Gln Arg Ala
                305                 310
```

<210> 66
<211> 135
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2274523CD1

<400> 66

```
Met Phe Ala Gln Pro Phe Ser Pro Ile Arg Ala Ser Lys Arg Met
  1                 5                  10                  15
Ala Lys Val Ser Ser Asn Asn Phe Ala Ser Leu Pro Arg Gln Ala
                 20                  25                  30
Pro Met Leu Leu Phe Cys Pro Leu Trp Met Pro Val Thr Ser Val
                 35                  40                  45
Pro Gln Glu Ala Lys Leu Leu Arg Gln Leu Lys Phe Ser Gln Gly
                 50                  55                  60
Thr Gly Val Cys Val Leu Ile Tyr Thr Pro Leu His Thr Tyr Phe
                 65                  70                  75
Phe Lys Leu Ser Pro Thr Leu Gly Thr Pro Val Leu Glu Tyr Pro
                 80                  85                  90
Gln Gln Phe Lys Gly Lys Lys Arg Leu Lys Gln Lys Asp Phe Phe
                 95                 100                 105
Leu Pro Lys Leu Cys Leu Leu Ala Trp Gly Pro Arg His Ala Asp
                110                 115                 120
Leu Lys Ile Asn Gln Ala Trp Val Gly His Gly Gly Ser Arg Leu
                125                 130                 135
```

<210> 67
<211> 205
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1801820CD1

<400> 67

```
Met Val Asn Leu Ala Ala Met Val Trp Arg Arg Leu Leu Arg Lys
  1                 5                  10                  15
```

```
Arg Trp Val Leu Ala Leu Val Phe Gly Leu Ser Leu Val Tyr Phe
                20                  25                  30
Leu Ser Ser Thr Phe Lys Gln Glu Glu Arg Ala Val Arg Asp Arg
                35                  40                  45
Asn Leu Leu Gln Val His Asp His Asn Gln Pro Ile Pro Trp Lys
                50                  55                  60
Val Gln Phe Asn Leu Gly Asn Ser Ser Arg Pro Ser Asn Gln Cys
                65                  70                  75
Arg Asn Ser Ile Gln Gly Lys His Leu Ile Thr Asp Glu Leu Gly
                80                  85                  90
Tyr Val Cys Glu Arg Lys Asp Leu Leu Val Asn Gly Cys Cys Asn
                95                  100                 105
Val Asn Val Pro Ser Thr Lys Gln Tyr Cys Cys Asp Gly Cys Trp
                110                 115                 120
Pro Asn Gly Cys Cys Ser Ala Tyr Glu Tyr Cys Val Ser Cys Cys
                125                 130                 135
Leu Gln Pro Asn Lys Gln Leu Leu Leu Glu Leu Phe Leu Asn Arg
                140                 145                 150
Ala Ala Val Ala Phe Gln Asn Leu Phe Met Ala Val Glu Asp His
                155                 160                 165
Phe Glu Leu Cys Leu Ala Lys Cys Arg Thr Ser Ser Gln Ser Val
                170                 175                 180
Gln His Glu Asn Thr Tyr Arg Asp Pro Ile Ala Lys Tyr Cys Tyr
                185                 190                 195
Gly Glu Ser Pro Pro Glu Leu Phe Pro Ala
                200                 205
```

<210> 68
<211> 2569
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3211795CB1

<400> 68

```
ctcggacccc tttaaggcgc ggccagagtc ctcccgcaga aaaacgactt aaaggagacg 60
cgtggcgcga tggcggcggc cccacgcgcg ggccggcggc gcgggcagcc gctcctggcg 120
ctgctgcttc tgctgctggc gccactgccg ccgggggccc cgccgggcgc cgacgcctac 180
tttcccgagg agcgctggag cccggagtcg cccctgcagg cgccgcgcgt gctcatcgcg 240
ctgttggcgc gaaacgcggc ccacgcgttg cccaccacgc tgggcgcact cgagcggctg 300
cggcacccgc gggagcgcac ggcgctatgg gtggctacgg accacaacat ggataacacg 360
tcaactgtgc tgcgggagtg gctggtggcc gtgaagagtt tgtaccattc cgtggagtgg 420
cggccagcag aggagcccag gtcctacccg gacgaggaag gcccgaaaca ctggtctgac 480
tcacgctacg agcatgtcat gaagttgcgc caggcagccc tgaaatcagc tcgagacatg 540
tgggctgatt acatcctgtt tgtagatgcg gacaacctga tcctcaaccc tgacacactg 600
agcctgctca tcgctgagaa caagacggtg gtcgccccca tgctggattc ccgggctgcg 660
tactccaact tctggtgtgg aatgacttcc cagggctact acaagcgcac acctgcctac 720
atccctatcc gcaagcgaga ccgccggggc tgctttgcag ttcccatggt gcactcgacc 780
ttcctgatcg acctgcggaa ggcggcgtcc aggaacctgg ccttctaccc acctcaccct 840
gactacacct ggtcctttga cgacatcatc gtctttgcct tctcctgcaa gcaggcagag 900
gttcagatgt atgtgtgcaa caaggaggag tacggattct tgccagtgcc attgcgcgcc 960
cacagcaccc tccaggatga ggccgagagc ttcatgcatg tgcagctgga ggtcatggtg 1020
aagcacccgc ccgcagagcc ctcccgcttc atctcggctc ccaccaagac accggacaag 1080
atgggcttcg acgaggtctt catgatcaac ctgaggcggc ggcaggaccg gcgggagcgc 1140
atgctgcggg cgctgcaggc acaggagatc gagtgccggc tggtggaggc cgtggacggc 1200
aaagccatga acaccagcca ggtggaggcg ctggggatcc agatgctgcc tggctaccgg 1260
gacccctacc acggccggcc cctcaccaag ggtgagctgg gctgcttcct gagccactac 1320
aacatctgga aggaggtggt ggaccggggg ctgcagaaat cgcttgtgtt tgaggatgac 1380
ctgcgttttg agatcttctt caagagacgt ctgatgaacc tcatgcggga tgtggagcgg 1440
gagggcctgg actgggacct catctatgtg ggccggaagc ggatgcaggt ggagcacccc 1500
gagaaggctg tgcctcgcgt gaggaacctg gtggaggccg actattccta ctggaccctg 1560
```

gcctacgtga tctccctgca aggcgcccgc aaactgctgg ctgctgagcc gctctccaag 1620
atgctgcctg tggacgagtt cctgcccgtc atgttcgaca aacacccagt gtccgagtac 1680
aaggcccact tctccctccg caacctgcat gccttctctg tggagccgct gctcatctac 1740
cccacacact acacaggaga cgatggctat gtgagtgaca ccgagacctc agtcgtatgg 1800
aacaatgagc acgtcaagac cgactgggac cgcgccaagt cccagaagat gcgggagcag 1860
caggcactga gccgtgaggc caagaactcg gacgtgctcc agtccccact ggacagtgct 1920
gcccgggatg aactctgagg ggtagcagcc agaaagccaa agcagccatc ggtggcccag 1980
gctccacgtg cttactgagg acatcagggc cacctctgga ccccttggca ggccacagag 2040
ggctctcgtg tggggtggtg tccagccagc tcttgctaag caatcacgtg cacacaggca 2100
gcattaatgg agtgcctact gcatgccagc aacagggctt ggccctgggg aattgggagg 2160
aaccaagccc tcttcatctg ttcatgtgcc cagcatttat taagcacctg ctgtatgcaa 2220
ggttcccatg ttacggcagt gaatgaggca taattgttcc ctccatcagc gattgattca 2280
gtcatcaagc agttactgat cagattaaga atcaggcact agtgatacac attcattttt 2340
aaaattcatt caaggattta ttgagtgcct actgtgtgtt gggtgccatt ccaggctctg 2400
ggatttttt ttttttttt ttttaagagt agagtctgtc tctgtcaccc aggctggagt 2460
gcagtggtgt gacggctcac tgcagcctgc gcctcccagc gtccagcaat tcttgtttct 2520
cggcctccca agtagctggg actataggtg cgtgccatca catctggct 2569

<210> 69
<211> 2387
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6813464CB1

<400> 69
atggcgggcg cccctccccc cgccttgctg ctgccttgca gtttgatctc agactgctgt 60
gctagcaatc agcgacactc cgtgggcgta ggaccctccg agctagtcaa gaagcaaatt 120
gagttgaagt ctcgaggtgt gaagctgatg cccagcaaag acaacagcca gaagacgtct 180
gtgttaactc aggttggtgt gtcccaagga cataatatgt gtccagaccc tggcataccc 240
gaaaggggca aaagactagg ctcggatttc aggttaggat ccagcgtcca gttcacctgc 300
aacgagggct atgacctgca agggtccaag cggatcacct gtatgaaagt gagcgacatg 360
tttgcggcct ggagcgacca caggccagtc tgccgagccc gcatgtgtga tgcccacctt 420
cgaggcccct cgggcatcat cacctccccc aatttcccca ttcagtatga caacaatgca 480
cactgtgtgt ggatcatcac agcactcaac ccctccaagg tgatcaagct cgcctttgag 540
gagtttgatt tggagagggg ctatgacacc ctgacggtcg gtgatggtgg tcaggatggg 600
gaccagaaga cagttctcta catcctgaca ggtacatcgg tcccggatct cattgtcagc 660
accaatcatc aaatgtggct cctcttccag actgatggca gtggcagttc cctgggattc 720
aaggcttctt atgaagagat cgagcagggc agttgcggtg accctggcat acctgcatat 780
ggccggaggg aaggctcccg gtttcaccac ggtgacacac tcaagtttga gtgccagccc 840
gcctttgagc tggtgggaca gaaggcaatc acatgccaaa agaataacca atggtcggct 900
aagaagccag gctgcgtgtt ctcctgcttc ttcaacttca ccagcccgtc tggggttgtc 960
ctgtctccca actacccaga ggactatggc aaccacctcc actgtgtctg gctcatcctg 1020
gccaggcctg agagccgcat ccacctggcc ttcaacgaca ttgacgtgga gcctcagttt 1080
gatttcctgg tcatcaagga tggggccacc gccgaggcgc ccgtcctggg caccttctca 1140
ggaaaccagc ttccctcctc catcacaagc agtggccacg tggcccgtct cgagttccag 1200
actgaccact ccacagggaa gaggggcttc aacatcactt ttaccacctt ccgacacaac 1260
gagtgcccgg atcctggcgt tccagtaaat ggcaaacggt ttggggacag cctccagctg 1320
ggcagctcca tctccttcct ctgtgatgaa ggcttccttg ggactcaggg ctcagagacc 1380
atcacctgcg tcctgaagga gggcagcgtg gtctggaaca gcgctgtgct gcggtgtgaa 1440
gctccctgtg gtggtcacct gacttcgccc agcggcacca tcctctctcc gggctggcct 1500
ggcttctaca aggatgcctt gagctgtgcc tgggtgattg aggcccagcc aggctacccc 1560
atcaaaatca ccttcgacag gtgcttgtga ccacagcctt gggggagagg gcaaggggga 1620
gatttttaca gagccacgtg ggagggaatg cagggtggat aggggaccag taccaacgct 1680
gaaatgtcag caaggttgga gaacactgga ccaaattgcc ctgggtgggg cagacctcac 1740
tggctctgcc cttgggcttg gtttgtgctg tttaggggac tatggcagga ggggtgaaga 1800
gaaggcttgg gtaccatttg ggatgttaga gctccccccc ccccagcccc caaggagggg 1860
gagtgggcag atagacacca tcagatgtta aagcatgact ctggtatctg acgagcattc 1920
tcctggaaat gtccaacgtg ggaaggcagc cccccacctc tgccccaaat cctgctcatg 1980
tgccccactt ctagctccca ccctggctat gccccttggg tggagaacgt agccacctaa 2040
atgtgaaaac aaacttcttc cggacaaagt gttcctggcc tggaatagca gtctgcaaac 2100

ttttgtgtga agagccagat ggtaaatatt ttaggctttg tgggccactt gcctacttgc 2160
atatctctga tgcatgctct gctcttggcg ccttttttt ttcttaagat ctctttacaa 2220
atgtaaaaag catccttagc tctccaggct atagaccaga tctgggccct gaggctctta 2280
gtgtgctgac ccctagccta gaatgttctg acctccatga ggctagtcat tgtgcctcta 2340
ggtcttaagc tgggctcaca gatcaggcca cagagaggat gagagct 2387

<210> 70
<211> 1959
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2156540CB1

<400> 70
ggccggctgg gcgtgcgctc gctccccgaa gccggggctg ggccggagcc gggcgagggc 60
tgggagctgg gccgggtccg gggacagcgg gcgaggggca gctgccggag ccgggcagcc 120
aggccgctca gggcagggga cagctggcgc cggttctgcg gtctccgggg cccagatgtg 180
aggcggcggc gccccggcc cgagagcgca cgatgggggc cccgctcgcc gtagcgctgg 240
gcgccctcca ctacctggca cttttcctgc aactcggcgg cgccacgcgg cccgccggcc 300
acgcgccctg ggacaaccac gtctccggcc acgccctgtt cacagagaca ccccatgaca 360
tgacagcacg gacgggcgag gacgtggaga tggcctgctc cttccgcggc agcggctccc 420
cctcctactc gctggagatc cagtggtggt atgtacggag ccaccgggac tggaccgaca 480
agcaggcgtg ggcctcgaac cagctaaaag catctcagca ggaagacgca gggaaggagg 540
caaccaaaat aagtgtggtc aaggtggtgg gcagcaacat ctcccacaag ctgcgcctgt 600
cccgggtgaa gcccacggac gaaggcacct acgagtgccg cgtcatcgac ttcagcgacg 660
gcaaggcccg gcaccacaag gtcaaggcct acctgcgggt gcagccaggg gagaactccg 720
tcctgcatct gcccgaagcc cctcccgccg cgcccgcccc gccgcccccc aagccaggca 780
aggagctgag gaagcgctcg gtggaccagg aggcctgcag cctctagact gatgcccctg 840
cccccgccca tccgccccca cgctgtacag agtgcatgag gagccgccgg accaccgggg 900
accgactgcc tgcgtccagc cgtgccccat ccccgaggcc gcctgtggcc accatgtcgg 960
ccctctttcc accacccctt gctcagcatg taagccccac ccacccctgc cctttcagac 1020
ccctgcggtg acctggctcg gagaaggtgg ccctgggcac caaggggcca accgccctga 1080
acactggggc agggaccatg ctgggcccg gggccacccc cttcctgtca ccagcttctg 1140
tggagtccag tgttttgctt tgcttgcttg tcccccatcc tgtcctgagc cggggccccc 1200
cagcctcgcc tccctcctcc taccatccct cacttggacc tgggggtgtg gacagtgacc 1260
cctccctgaa tatggacttg aatcttctga gcagaactag ggcctctccc ctggtgaaga 1320
cccagggaac ccaggagggc ccttctgggg cagtggctct gcagggtcac tcatggaggc 1380
ctaggggaac agcgagatgc cccaccacct cctggcgagt ccttcctgtt cagctccctg 1440
tgcgaccctc cagggatgca ggggatccag gattctctgc cctgtcacac ggcgagtcag 1500
aagggagggg cctttccctc ggacccatgg ccccaggcag agttttgcac cagcaggacc 1560
cctttgaggg ccttcaaggc tctcccagga gtcccccctct gccggccccc caatgcccca 1620
gctccctctt gggtcctgtg ccaagtccgc cccagggcct ggggctgttg ggagccaagg 1680
gccccctggt actcagttcc ctcacgattc ccgatcacgg gcacacctgc cccctggtta 1740
tttgtaaata tttctattgg acccaattct cctcggaatt ggctggcacc tctggctgcc 1800
acagctcagt gatgacgtgg gggaggtggg agaggccgag ggctttgcct aggggtgggt 1860
tgccctgtat acatgatcca gtctgtgact accagccaac ctgaataaag cggttttaaa 1920
aaaaaaaaaa aaaaaaaaat atgcggcgca agcttattc 1959

<210> 71
<211> 1562
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 894939CB1

<400> 71
cctcttccgt cggctgaatt gcggccgtat gcgcggctct gtggagtgca cctggggttg 60
ggggcactgt gccccccagcc ccctgctcct ttggactcta cttctgtttg cagccccatt 120
tggcctgctg ggggagaaga cccgccaggt gtctctggag gtcatcccta actggctggg 180

```
ccccctgcag aacctgcttc atatacgggc agtgggcacc aattccacac tgcactatgt 240
gtggagcagc ctggggcctc tggcagtggt aatggtggcc accaacaccc cccacagcac 300
cctgagcgtc aactggagcc tcctgctatc ccctgagccc gatgggggcc tgatggtgct 360
ccctaaggac agcattcagt tttcttctgc ccttgttttt accaggctgc ttgagtttga 420
cagcaccaac gtgtccgata cggcagcaaa gcctttggga agaccatatc ctccatactc 480
cttggccgat ttctcttgga acaacatcac tgattcattg gatcctgcca ccctgagtgc 540
cacatttcaa ggccacccca tgaacgaccc taccaggact tttgccaatg gcagcctggc 600
cttcagggtc caggcctttt ccaggtccag ccgaccagcc caacccctc gcctcctgca 660
cacagcagac acctgtcagc tagaggtggc cctgattgga gcctctcccc ggggaaaccg 720
ttccctgttt gggctggagg tagccacatt gggccagggc cctgactgcc cctcaatgca 780
ggagcagcac tccatcgacg atgaatatgc accggccgtc ttccagttgg accagctact 840
gtggggctcc ctcccatcag gctttgcaca gtggcgacca gtggcttact cccagaagcc 900
gggggggccga gaatcagccc tgccctgcca agcttcccct cttcatcctg ccttagcata 960
ctctcttccc cagtcaccca ttgtccgagc cttctttggg tcccagaata acttctgtgc 1020
cttcaatctg acgttcgggg cttccacagg ccctggctat tgggaccaac actacctcag 1080
ctggtcgatg ctcctgggtg tgggcttccc tccagtggac ggcttgtccc cactagtcct 1140
gggcatcatg gcagtggccc tgggtgcccc agggctcatg ctgctagggg gcggcttggt 1200
tctgctgctg caccacaaga agtactcaga gtaccagtcc ataaattaag gcccgctctc 1260
tggagggaag gacattactg aacctgtctt gctgtgcctc gaaactctgg aggttggagc 1320
atcaagttcc agccggcccc ttcactcccc catcttgctt ttctgtggaa cctcagaggc 1380
cagcctcgac ttcctggaga cccccaggtg gggcttcctt catactttgt tgggggactt 1440
tggaggcggg caggggacag ggctattgat aaggtccccc tggtgttgcc ttcttgcatc 1500
tccacacatt tcccttggat gggacttgca ggcctaaatg agaggcattc tgactggttg 1560
gc                                                                1562

<210> 72
<211> 3425
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4620890CB1

<400> 72
tcccagggtg atcccgagca cggccggtgg ctcgcggcgc ggcagcccca gaagacggga 60
aagttcgcgg ccggagcgcg gagatgccgg gcagcgacac ggcgctcacc gtggaccgga 120
cctactcgga ccccggccgg caccaccgct gcaagagccg ggtagaacgt catgacatga 180
ataccttaag cctgcccctg aacatacgcc gaggggggtc agacaccaac ctcaactttg 240
atgtcccgga tggcatcctg gacttccaca aggtcaaact cactgcagac agcctgaagc 300
aaaaaattct aaaggtaaca gagcagataa aaattgagca aacatcgcgc gatgggaatg 360
ttgcggagta tctgaaacta gtgaacaacg cggacaagca gcaggcggga cgtatcaagc 420
aagtctttga gaagaagaat cagaaatcag ctcactccat cgcccagctg cagaagaagt 480
tagagcagta tcatcgaaag ctcagagaga tcgagcagaa tggagcctct aggagctcaa 540
aggacatttc caaagaccac ctgaaggata tacatcgctc tttgaaagat gcccacgtga 600
aatctcgaac tgccccccat tgcatggaga gcagcaaatc gggcatgcca ggggtctcac 660
ttactccacc tgtgtttgtt ttcaataagt ccagagagtt tgccaacctg atccggaata 720
agtttggcag cgccgacaac attgctcact tgaaaaattc cttagaggag tttaggccag 780
aggcgagtgc cagggcctac gggggcagcg ctaccatcgt gaacaaaccc aagtatggca 840
gtgatgatga atgttcgagt ggcacgtcag gctcggccga cagtaacgga aaccagtcgt 900
ttggggctgg tggagccagc acactggaca gccagggcaa gctcgccgtg atcctggagg 960
aactgaggga gatcaaggat acccaagctc agctggctga ggacatcgag gcactgaagg 1020
tgcagtttaa gagagaatat ggttttattt ctcagaccct gcaagaggaa agatacaggt 1080
atgagcgact ggaagaccag ctgcatgacc tgacggacct gcatcagcat gagacagcca 1140
acctgaagca ggagctggcc agcattgagg agaaggtggc ctaccaggcc tacgagcgct 1200
cgcgggacat ccaggaagcc ttggaatcct gccagactcg catttctaag ctggagctcc 1260
accagcaaga gcagcaagct ctgcagacag acaccgtgaa tgctaaagtt ctcctgggga 1320
ggtgcatcaa cgtgatcctg gccttcatga ctgtcatctt agtgtgtgtg tccaccatcg 1380
cgaagttcgt ctcacccatg atgaagagtc gctgccacat tcttggcacc ttctttgccg 1440
tgactcttct tgctatattt tgtaaaaact gggaccatat cctgtgtgcc atagaaagga 1500
tgataatacc aagatgaagc cactggttcc tgccttcaag ttctttcaag tttttatttt 1560
aaagaaaact ctgtgcatac taccaaattt tacagtgaat gattgtgcgg actcgtgtgt 1620
aagaaaaact aggactgtgt ggtgtaaata actacaattc tcttaactcg gtagcagttg 1680
```

ccaactcagt ccttgtactt cgttaacacg aatctgtttc agagctctcc taccttgctc 1740
actgccttaa tcagaccgat ttcctgccca cctgaccagc ccagcgtggt aaacctctgt 1800
atattgagac cttggcataa ttggtgatcc tgaagaaaga ggtctctctc ctaagtctct 1860
gtcagaattg agcttcacaa ttgctaatgg ttgttttctg tgagtcctat aaaaagcaag 1920
gatatgcatg attcagggaa tgaagaatca caggcttggg cagtgttaaa cactgtggcc 1980
tatggtcccc gtgtgatcca ccctgcttct ctccagggga ccataggtcc cgtcatgtac 2040
tcagtgtcca cagcagtcag tcgtgtatga ccctgtaacg tggaaatctt atcacacacc 2100
tgttatccaa caagtctacc tgaggggttt tgttacactt taaatgggaa ggcataggga 2160
tttatgaatg gggctttcac cttctcatac ccaggcaacc aacacctgat tttgtctcaa 2220
ctggctagca aatgcccagc cttcagagtg tgcaggaatg ttttcaaatc cctcatcaga 2280
ctgtgacttt aacattaatt tggaatcctg tgagcactac tctgaaggtt tgtgttttgg 2340
caaatctttt ttctttttg agacagggct ctgctaaata ttgctcaggc tggttgcaaa 2400
ctccttgctt caagggatcc tcccacctca gcctcccaag cagccgggac tgcaggcaca 2460
agccaccatg cctggctgtt ttttggcaaa tcttgattgt gataagcccc cctggaggat 2520
atgattcact ttatgtgatt catcttattc acaggtctgt gagggactgc gaagcttact 2580
caggaaatga aaacaaatga tggtcatgtt gcagtttttt ccttgaagga caaccgaacc 2640
atagcctcta aagttcaagt gcactgaggt gtcggaacgc tgaaagcatg aggaaacgag 2700
gacgtagggt gtgactgaat ggtggctaga ttagtgggag cagttcacct ggatgaagat 2760
tgagagcatc gtctttgaga agtgaaagac tagcaagaat aaaataaatt aagtccagtg 2820
tttgagccaa ggttgccacc tgtctcttaa catctcactg aacataagtc ctgaggtatt 2880
aggacgacca tactgcctct gagctgaaaa cattcaaaag ttcacatccc tgtttggggg 2940
ataccattca ccgccttcag cccagatgat actttccttt aaatctgtgt ctctgtgtgt 3000
ataacaaaga ggaagatgga aacaatgttc atggaaactg ctgttgagcc ccttgtccca 3060
ccactcccgc catctgctgc aggcaggaag gcatgtgagt gtacgttttc ttccaggaga 3120
catcaggtcc ccctggattc aaattaagtg caatattttg caaacagctc ttcttaggga 3180
aatctcctga aggaaaaaaa tgtgacagaa tgttccatag tctgagagaa tggaatcgtt 3240
gagcatttag tacaagtcca gtgtgtgtga gcgggactta ggcagctcaa gcttgctttt 3300
ttttttaagc gtacaattga gtggttttag taaattcaca aacttgttca accatcacca 3360
ctatctaatt ccagactcac gctttttttaa acaataaatg tcatttcatg aaaaaaaaaa 3420
aaaaa 3425

<210> 73
<211> 3130
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5514146CB1

<400> 73
gggacagtca aagaatttta aaaacaggca actttgtaac tgtgaaatac tctccaggat 60
ttaaaaggct gtggagctcc agataaagaa tcgtttatct ttcttctgaa gaaattcctt 120
tggttacaag tttaccccat aaacggcaac acactcacct ccatccaaga cagactcaag 180
gtggaggaag cgtggaaatg tgcttccgga caaagctctc agtatcctgg gtgccattgt 240
ttcttctact cagccgtgtt ttttctactg agacagacaa accctcagcc caggacagca 300
gaagccgtgg gagttcaggc caaccggcag acctgctaca ggttctctct gctggtgacc 360
acccacccca caaccactca agaagcctca tcaaaacatt gttggagaaa actgggtgcc 420
cacggaggag aaacggaatg caaggagatt gcaatctgtg ctttgaacca gatgcactat 480
tactaatagc tggaggaaat tttgaagatc agcttagaga agaagtggtc cagagagttt 540
ctcttctcct tctctattac attattcatc aggaagagat ctgttcttca aagctcaaca 600
tgagtaataa agagtataaa ttttacctac acagcctact gagcctcagg caggatgaag 660
attcctcttt cctttcacag aatgagacag aagatatctt ggctttcacc aggcagtact 720
ttgacacttc tcaaagccag tgtatggaaa ccaaaacgct gcagaaaaaa tctggaatag 780
tgagcagtga aggtgctaat gaaagtacgc ttcctcagtt ggcagccatg atcattactt 840
tgtccctcca gggtgtttgt ctgggacaag gaaacttgcc ttccccagac tactttacag 900
aatatatttt cagttccttg aatcgtacga atacactccg cctatcagaa ctagaccaac 960
tcctcaacac tctctggacc agaagtactt gtatcaaaaa tgagaaaatc catcaatttc 1020
aaaggaaaca aaacaacata ataacccatg atcaggacta ttctaatttc tcttcatcca 1080
tggaaaaaga gtctgaggat ggtccaattt cctgggatca gacctgcttc tctgctaggc 1140
agctggtgga gatatttcta cagaagggcc tctcactcat ttctaaggag gactttaagc 1200
aaatgagtcc agggatcatc cagcagctcc tcagctgctc ctgccactta cccaaggacc 1260
aacaagcaaa gctgccacct accactctgg agaaatacgg ctacagcacg gtggctgtca 1320

ccctttctcac actgggctcc atgctgggga cagcgctggt ccttttccat agctgtgagg 1380
agaactacag gcttatctta cagctgtttg tgggcttggc cgtcgggaca ctgtctgggg 1440
acgctctgct ccaccttatc cctcaggttc ttggtttaca taagcaggaa gccccagaat 1500
ttgggcattt ccatgaaagc aaaggtcata tttggaaact gatgggatta attggaggca 1560
tccatggatt tttcttgata gaaaaatgtt ttattcttct tgtatcacca aatgacaagc 1620
agggcctgtc attggttaat gggcacgtgg gtcattccca ccatcttgca ctcaactctg 1680
aattaagtga ccaggcaggc agaggcaaat ctgcttcaac tatccagttg aaaagcccag 1740
aagattcaca ggcagctgaa atgcctatag gcagtatgac agcctccaac agaaaatgta 1800
aagccattag cttgttagca atcatgattc tggttgggga cagcctgcat aattttgcag 1860
atggcctagc cataggagca gccttctcat catcatccga gtcaggagtg accactacga 1920
ttgctatctt gtgtcatgaa atcccacatg aaatgggaga ctttgccgtg ctcttaagct 1980
ctggactttc tatgaagact gccatcctga tgaattttat aagctcccta actgccttca 2040
tgggattata cattggcctt tccgtgtcag ctgatccatg tgttcaagac tggatcttca 2100
cagtcactgc tgggatgttc ttatatttat ccttggttga aatgcttcct gaaatgactc 2160
atgttcaaac acaacgaccc tggatgatgt ttctcctgca aaactttgga ttgatcctag 2220
gttggctttc tctcctgctc ttggctatat atgagcaaaa tattaaaata taagtgagga 2280
tcttcaacat ctttcaaaaa tgcatttata tagtcttact ttgtttcttt cattgcactc 2340
tataatgatt tttaaattaa gaatttttta tcttaggcaa agtgtgtctc tttcaattca 2400
ttaacttatt aattttataa tgcagtttta tttttggaaa catataaata tcagactgtc 2460
cttaattgaa attttgtctt tggtttccaa caccatgatg aagctcttgc tttttaaaaa 2520
gtagttagta aattctgcat gaatttttagt aaactttaaa aaatagattt tttccctaag 2580
aaagaatgtt tgtagaattt aaagtggaca gatgcctgtt ggggtaaaat caactgcaac 2640
tttttgatgt taatttttttt ccctgtgcaa ttataaacta taagcaagtt aagtgacaag 2700
caaatgtaat aaagactagt tttaaaaaaaa aaaaaaaggg gggttttctc cccggccctg 2760
ggctatattt gggcaaaatt ttaaaatata agtggggttc ttcaacattt ttcaaaaagc 2820
cattaatagg gtctaacttg gttccttcca tggcactcta aaaggttttt aaatttagga 2880
ttttttttacc ttgggcaaag tggggctctt ccaatcctta acctttttaaa ttttataagg 2940
cggttttatg tttggaaaac cttttaaatt ccgagcgggc ctaaatggaa atttggcttt 3000
gggttcccaa ccccgtggtg aacccctttgc ttttaaaaag tagtgggggc ccgggcccga 3060
atccgaaaca tgctaaaggg gttccccggg gaatttttta cccgccaaaa ttccccaaaa 3120
tatggggcgg 3130

<210> 74
<211> 3172
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7474769CB1

<400> 74
ggctcctgta gacaatatac ccgtgaatgg tgtagtgagt tctaatgaaa ctttatttac 60
aagaggagac tgaccaggtt tggcctgggg gccacagtgt gtagacctct ggaaagatac 120
atcctgagaa gaaaaaaaga atatatgcag gaatgtttaa ctttgtgggt tttctctcct 180
cttgccctca ctgactcagg atacacaaag acctatcaag ctcacgcaaa gcagaaattc 240
agccgcttat ggtccagcaa gtctgtcact gagattcacc tatactttga ggaggaagtc 300
aagcaagaag aatgtgacca tttggaccgc ctttttgctc ccaaggaagc tgggaaacag 360
ccacgtacag tgatcattca aggaccacaa ggaattggaa aaacgacact cctgatgaag 420
ctgatgatgg cctggtcgga caacaagatc tttcgggata ggttcctgta cacgttctat 480
ttctgctgca gagaactgag ggagttgccg ccaacgagtt tggctgactt gatttccaga 540
gagtggcctg accccgctgc tcctataaca gagatcgtgt ctcaaccgga gagactcttg 600
ttcgtcatcg acagcttcga agagctgcag ggcggcttga acgaacccga ttcggatctg 660
tgtggtgact tgatggagaa acggccggtg caggtgcttc tgagcagttt gctgaggaag 720
aagatgctcc cggaggcctc cctgctcatc gctatcaaac ccgtgtgccc gaaggagctc 780
cgggatcagg tgacgatctc agaaatctac cagccccggg gattcaacga gagtgatagg 840
ttagtgtatt tctgctgttt cttcaaagac ccgaaaagag ccatggaagc cttcaatctt 900
gtaagagaaa gtgaacagct gttttccata tgccaaatcc cgctcctctg ctggatcctg 960
tgtaccagtc tgaagcaaga gatgcagaaa ggaaaagacc tggccctgac ctgccagagc 1020
actacctctg tgtactcctc tttcgtcttt aacctgttca cacctgaggg tgccgagggc 1080
ccgactccgc aaacccagca ccagctgaag gccctgtgct ccctggctgc agagggtatg 1140
tggacagaca catttgagtt ttgtgaagac gacctccgga gaaatggggt tgttgacgct 1200
gacatccctg cgctgctggg caccaagata cttctgaagt acggggagcg tgagagctcc 1260

tacgtgttcc tccacgtgtg tatccaggag ttctgtgccg ccttgttcta tttgctcaag 1320
agccaccttg atcatcctca cccagctgtg agatgtgtac aggaattgct agttgccaat 1380
tttgaaaaag caaggagagc acattggatt tttttggggt gttttctaac tggcctttta 1440
aataaaaagg aacaagaaaa actggatgcg ttttttggct tccaactgtc ccaagagata 1500
aagcagcaaa ttcaccagtg cctgaagagc ttaggggagc gtggcaatcc tcagggacag 1560
gtggattcct tggcgatatt ttactgtctc tttgaaatgc aggatcctgc ctttgtgaag 1620
caggcagtga acctcctcca agaagctaac tttcatatta ttgacaacgt ggacttggtg 1680
gttttctgcct actgcttaaa atactgctcc agcttgagga aactctgttt ttccgttcaa 1740
aatgtcttta agaaagagga tgaacacagc tctacgtcgg attacagcct catctgttgg 1800
catcacatct gctctgtgct caccaccagc gggcacctca gagagctcca ggtgcaggac 1860
agcaccctca gcgagtcgac ctttgtgacc tggtgtaacc agctgaggca tcccagctgt 1920
cgccttcaga agcttggaat aaataacgtt tccttttctg gccagagtgt tctgctcttt 1980
gaggtgctct tttatcagcc agacttgaaa tacctgagct tcaccctcac gaaactctct 2040
cgtgatgaca tcaggtccct ctgtgatgcc ttgaactacc cagcaggcaa cgtcaaagag 2100
ctagcgctgg taaattgtca cctctcaccc attgattgtg aagtccttgc tggccttcta 2160
accaacaaca agaagctgac gtatctgaat gtatcctgca accagttaga cacaggcgtg 2220
ccccttttgt gtgaagccct gtgcagccca gacacggtcc tggtatacct gatgttggct 2280
ttctgccacc tcagcgagca gtgctgcgaa tacatctctg aaatgcttct gcgtaacaag 2340
agcgtgcgct atctagacct cagtgccaat gtcctgaagg acgaaggact gaaaactctc 2400
tgcgaggcct tgaaacatcc ggactgctgc ctggattcac tgtgtttggt aaaatgtttt 2460
atcactgctg ctggctgtga agacctcgcc tctgctctca tcagcaatca aaacctgaag 2520
attctgcaaa ttgggtgcaa tgaaatcgga gatgtgggtg tgcagctgtt gtgtcgggct 2580
ctgacgcata cggattgccg cttagagatt cttgggttgg aagaatgtgg gttaacgagc 2640
acctgctgta aggatctcgc gtctgttctc acctgcagta agaccctgca gcagctcaac 2700
ctgaccttga acaccttgga ccacacaggg gtggttgtac tctgtgaggc cctgagacac 2760
ccagagtgtg ccctgcaggt gctcgggctg agaaaaactg attttgatga ggaaacccag 2820
gcacttctga cggctgagga agagagaaat cctaacctga ccatcacaga cgactgtgac 2880
acaatcacaa gggtagagat ctgattgcga ggaacctggg ctctgactcg aacacctgca 2940
aaggacaggg actgggaccg ttacttacat gacactgcac ccaggagata caaatcattg 3000
acactctgag ttgtgagatt tctggcaccc cattcataga tttgatatga tacacgtggt 3060
ttttatgtgc tctgtggcct tggatgagtc actgaaaggc cttcatggtc tctcggtctc 3120
acaaggacct cttaacccct caataaagtg ttacatttct aaacattgga aa 3172

<210> 75
<211> 2094
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 065296CB1

<400> 75
tgtgtgctga ttttttaatg aacaacttat aatcttacaa aataactggc tgtagctgtt 60
ttacattaca atacagtaga tcagctcctt tcactacatt cccaatcaag ccatcccctt 120
gggtgttaac taatgcactc aatagtgatc acaaagaata caggaagtgc caagttcaat 180
gccttgtaat agaaaaaaga gaagatgaga atcaggaaga gctctttact ttcacaatcc 240
ctttacttta atttctgtca gtatctgccc attctcctca ccctgtcccc ttacctgcct 300
ttctgattgg aggctgtcat gctagttgtt agagccagct agccctgggc cacctgggca 360
caatcaaaca cacagaagag tctctgagaa aggctctcaa tgaccacatg ggtggataag 420
agtagcagaa tgtggaccac cacatggatt gagtgtgagt tgatgttgct gggatctgtg 480
accacagtgt cttagtcctt ccactttatc agtgacctga gttgctcaaa acatcaggtc 540
ccaggaacaa agactggaag aacctgctgt ctctgaatag ctatagctga ctatggctgg 600
gaagccttat agcctaagag gctcttcaca cacaacagga actttccttt tgctttccca 660
gagcagtggt gaattgcaga taataaaata ttttaaaatg aaatttaaaa cagaaatgtt 720
cttactgttg ctcttgtgga gagactgcat gaagacacac actggcatga accacaggct 780
ccatgttcct gagctgagta atgcccagga taacaattcc tctgcctcaa tctcagacaa 840
agtaggattc tccaaggcag aacttagaat gtgtctagca atttggacat tttccccaat 900
aaaacaagtt tataaaattt tgaaaataga gtgtctaaac ttttccattg tccttagcgt 960
attgaagcct atcagaatcc caagaataaa tatgtttgtc ttcttggggg cactaagcat 1020
gacccaagac aatgaatggt atctgaacta cattttcttt acactagaga ttagtagaca 1080
aaaggttttt tttgaatggg tgaactctgc tctttccttt tcccaataac tttgcaaccc 1140
tcattccaca aacccacaga ttaatgtggg tacatcacct tattcgtgtt ttattctctt 1200

tcaggatatt gttcttccct acgtgatagt gtggataata tttctatcat tgtggtagta 1260
tttcctattc ctatttgcac agccattgct ctagtgttga tgggaagagt agaaaaactg 1320
tggtcaaagc aaaagagcag aagatgagaa gagaccacag ctgagctcta aattttgtct 1380
gttcttggtt tccattggag aagaaacttt aaaggagaca agcctcactc acagccttcc 1440
taaaatgagt aaagctagaa tgttaaacag aaactggctc cttactcact ttccctaggg 1500
tctaatccag aaaccgttac aagaaaatgt aactgcagta gaccacatcc cagtgagcct 1560
gggatgtaaa tagagagagg cttaatatgg tcatttcatg tgcattccat ctattggata 1620
agaaagggta aaaacaacaa aatggacatt aaatagttga aattcttctg tctccccaca 1680
gagtgtcacc cttgctccat gcatttccat ggcttgctca agtcaaaatc tctggctcac 1740
agattttgta catgatactt gaagagaatt tatctcagat tcttcattgc tcaagggacc 1800
cactttggtc ttctttggag acttggaaca gaggaccaag tctccctgat cacttccatt 1860
ctcatgcccg ggccctggat ttaacccctg acaggctctt ggcctgttct ttcatctgct 1920
tccagctttt ccctattcat ttccatgaaa ttttcctcca gaaaattaaa gttggtctgt 1980
gactggcgac ttgggacttc tacagaagtt ggatcgggca gtttgtggtt gtgtggtcct 2040
ggatcggggt caaaggggtg tcatttgtgg ggcccctggt ggtttttgga accg 2094

<210> 76
<211> 1119
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 231994CB1

<400> 76
ggcggagctg agggacccgg cggctggtga gcgcccgctg gaggctggag cttccgggcc 60
ctggaaaggg ggtccccgcg cgccccgggt cggaggcaga cccctgggtt tgggggacat 120
gggcatttgg ggcgcctgaa cccaagacct ctggatgagc tgccccgttc agaccatgga 180
tcctgaggtg accttgctgc tgcagtgccc tggcgggggc ctgccccagg agcagataca 240
ggccgagctg agccccgccc atgaccgtcg cccactgcca ggtggggacg aggccatcac 300
tgccatctgg gagacccggc taaaggccca accctggctc ttcgacgccc ccaagttccg 360
cctgcactca gccaccctgg cgcctattgg ctctcggggg ccacagctgc tcctgcgcct 420
gggccttact tcctaccgag acttcctggg caccaactgg tccagctcag ctgcctggct 480
gcgacagcag ggtgccaccg actggggtga cacgcaggcc tatctggcgg acccactggg 540
ggtgggcgct gcactagcca cagccgatga cttccttgtc ttcctgcgcc gctcccggca 600
ggtggctgag gcccctgggc tggtggacgt acctggtggg caccctgagc ctcaggccct 660
gtgccctggt ggcagccccc agcaccagga cctcgctggg cagctggtgg tacatgaact 720
cttttccagt gtccttcagg agatctgtga tgaggtgaac ctgccgctgc tcaccctgag 780
ccagcccctg ctgttgggca tcgcccgaaa tgagaccagt gctggccgag ccagtgccga 840
gttctatgtc cagtgcagcc tgacttctga gcaggtgagg aagcactacc tgagtggggg 900
acccgaggcc cacgagtcta caggaatctt ctttgtggag acacagaacg tgcggagatt 960
gcccgagacg gagatgtggg ctgaactctg cccctcggcc aaaggcgcca tcatcctcta 1020
caaccgggtt cagggaagtc ccactggagc ggccctaggg tccccagccc tactcccgcc 1080
gctctgaaaa taataaacga ctttattctt ggaaaaaaa 1119

<210> 77
<211> 3321
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 538054CB1

<400> 77
gtcgccgggt gaggaagcgg ggggctagcg gcctggcgct gcggcgaccg accggggcgt 60
caggatccct ggcccccgga gagcgaaggg cgggcgggtc ccggagcaag aagaattgta 120
tagaaatgta aaaaacattt atgtgtgctt cacatgtggc atgaaaaaaa tgaggacctc 180
tggtcaaaat tgcctgagtt cagaatttca tgctggaagg aaagttagac gtaacttcta 240
atttctagtt ccagtaagca tctcttcttg gaaaatttcc aaaaaagagt cctgctctgt 300
tgcccaggct ggagtacatt ggtgtgatga tcatggttca ctgcagcctc gacctcccag 360
gctcaagcaa ttctcctacc tcagcctcct gagtaactgg gactatagga tttctttgaa 420

gaaaaatgtg cttcttgaat caactgcagt ctctcctctg cataagagct cttaagggta 480
ctagcaggat agaagcaaat gaaactgaaa gctcatctgc agctcagaaa agcaaagaca 540
tggaatttta aagagtgaag gtagcatggt gtcggccatg ggtgaacaag acacagccag 600
acaatgtgga ccaatttctt caaactacgg cttttctgct gtctgcttgc agtgttgatg 660
gtggtggtgc tggtcatcaa tgttactcag gtagagtact tggaccatga gactgtttca 720
gccactttca tcgacagcag tggacagttt gtttcctccc aggtgacagg aattagccga 780
aatccctact gtggctatga tcagcagacc ctgtccagcc aggagcgcat ggaggaggac 840
tccttgctgg ctgccttgca ccggcaggtt cctgatgtgg gcccagtccc ctttgtgaag 900
agcactgacc cttcttccag ctactttgtc atcttgaact ctgctgcctt ctttaaggtg 960
ggaagccagc ttgaggtgct ggttcatgtg caggattttc aaagaaagcc caagaagtat 1020
ggtggagact acctgcaggc cagaattcac tccctcaagc tgcaggctgg ggctgtgggc 1080
agggtggtgg attaccagaa tgggttttac aaggttttct ttactttgct atggccaggc 1140
aaagttaaag tatccgtatc tctggtccac cccagtgaag ggatcagagt tcttcagcgc 1200
ttacaggaag ataaaccaga cagggtctat ttcaagagtc tcttccgttc aggaagaatt 1260
tctgaaacta ctgagtgcaa cgtgtgtctt cctgggaatc tgcccctgtg taactttaca 1320
gacctctaca ctggggagcc ctggttctgc ttcaaaccaa agaagctccc ttgcagcagc 1380
agaattaccc atttcaaagg tggatacctg aaaggtctcc taaccgctgc agagagtgct 1440
ttcttccaga gtggtgtcaa tatcaaaatg ccagtcaact ccagtggacc tgattgggta 1500
actgtgattc ccaggagaat aaaagaaact aacagtctag aactatctca aggctcagga 1560
acttttcctt ctgggtatta ttataaagac cagtggaggc ccagaaagtt taagatgcgt 1620
cagtttaatg accctgacaa cattacagag tgcttacaaa gaaaagtggt gcatttattt 1680
ggtgactcaa caatcaggca atggtttgaa taccttacta catttgttcc agatttagtg 1740
gagtttaact tgggtagtcc caagaatgtg ggtcccttcc ttgcagtgga ccagaagcac 1800
aacatcctgc tcaaataccg ctgccatggt ccacccatcc gcttcacgac tgtctttagc 1860
aatgagctcc attatgtggc gaatgagctg aatggcattg tgggagggaa gaacacagtg 1920
gttgccatag ctgtatggtc tcacttcagc accttccctt tggaagtgta catccggcgg 1980
ctcaggaaca tccgtcgagc agtggttcgg ctcctcgatc gaagcccaaa gaccgtggtg 2040
gtcatccgga cggccaacgc ccaagagctg ggacctgagg tgagcctttt caacagcgac 2100
tggtacaact ttcagctgga caccatcctt cggaggatgt tctcaggggt tggagtatat 2160
ctcgtcgatg cctgggagat gaccctggcc cattatctac cgcacaagct gcatccagat 2220
gaagttattg tgaagaacca gctggacatg ttcttgtcct ttgtgtgccc cttggaaacc 2280
tagcctgtct tggaagggac tggaggaatc atattcaatg accttctcaa ttgacctgag 2340
ttacagaaag tggccccagt gagagatgac tgcccttaat aagtataaaa tttcaaaaag 2400
atctggactt aatatgatga cttataagga gcttagaaaa tgcaggttac atttatatct 2460
acctatagga ttttatccaa tgttgactta gccatggtag aactcttaac tgcatctaca 2520
cactatattg ctcttgtaac caaagatgct aatgagtgta tttgaattag cttctcctag 2580
gaggggtgac tactttgcta aagagtatga aaaatgtttt gatggaaggg acaagtttgg 2640
ttggtagtag agtgtttgta gcttatctgg taaaggaaat tgagtgcatc tgcatgcata 2700
gcacaagtaa tccatcttct ggcggttaac tagagaaagc ttatagtcag aggctcttgt 2760
gaatgcactg aagactgctg tgtctcagga gtttctcttg atgatctacc ttttctgaat 2820
tacttgaaaa ggccaatgtg cttggttcga gtttttttagt gaatatttac aatttcctgc 2880
ttttatctag aaaagaagca aaagggaaat atgaaagcag tatatataaa atctgtgctt 2940
tggaattgat atttgcattt ggtgattgtt ttttatttga ttcagctaaa ttttatgtgt 3000
tgaatactta ctctagaata actacttttg gaactggaag ggggaaatct gtaaatggaa 3060
atttatttttt atgttctgaa atttgagggt ttaggaattg acttttgtgt aaagaatgct 3120
acggtattat aaagctgtgg aagtttttttc cactagtttc aggtgttttt ttttgttgtt 3180
gttgttgttg ttattttgaa ctacttacca aaattagaag ttaaaatcta attagaaatt 3240
aataagatac taaaattcat gacaaaaaaa aaaacaaaaa acacaaaaaa aaaaaaaaaa 3300
aaaaacaaaa aaaaaaaaaa a 3321

<210> 78
<211> 2646
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1259305CB1

<400> 78
gtcccttccc ctcacccgct ccacgccctc ctgggccgag tggagttggg tggtgtcggg 60
agcctctccc tgaggggcac cgcgtcttca ggagctgggc ctccagtgcg gcgcgatgtc 120
aggcgcggtg acagctctgt gagtccgagg ccgcggccgt ggcgctgggc ggctgcgggg 180

cctgaccggt ccgctcatgg tgccgccacg acgccatcgc ggggcaggaa ggccaggggt 240
gctgagttct tcacctcctt ttagactgag atctgccaag ttttccggca ttgctcttga 300
ggatctcaga agggctctta agacaagact gcaaatggtg tgtgtatttg tcatgaaccg 360
aatgaattcc cagaacagtg gtttcactca gcgcaggcga atggctcttg ggattgttat 420
tcttctgctt gttgatgtga tatgggttgc ttcctctgaa cttacttcgt atgttttac 480
ccagtacaac aaaccattct tcagcacctt tgcaaaaaca tctatgtttg ttttgtacct 540
tttgggcttt attatttgga agccatggag acaacagtgt acaagaggac ttcgcggaaa 600
gcatgctgct ttttttgcag atgctgaagg ttactttgct gcttgcacaa cagatacaac 660
tatgaatagt tctttgagtg aacctctgta tgtgcctgtg aaattccatg atcttccaag 720
tgaaaaacct gagagcacaa acattgatac tgaaaaaacc cccaaaaagt ctcgtgtgag 780
gttcagtaat atcatggaga ttcgacagct tccgtcaagt catgcattgg aagcaaagtt 840
gtctcgcatg tcatatcctg tgaaagaaca agaatccata ctgaaaactg tggggaaact 900
tactgcaact caagtagcga aaattagctt tttttttgc tttgtgtggt ttttggcaaa 960
tttgtcatat caagaagcac tttcagacac acaagttgct atagttaata ttttatcttc 1020
aacttccgga ctttttacct taatccttgc tgcagtattt ccaagtaaca gtggagatag 1080
atttaccctt tctaaactat tagctgtaat tttaagcatt ggaggcgttg tactggtaaa 1140
cctggcaggg tctgaaaaac ctgctggaag agacacagta ggttccattt ggtctcttgc 1200
tggagccatg ctctatgctg tctatattgt tatgattaag agaaaagtag atagagaaga 1260
caagttggat attccaatgt tctttggttt tgtaggtttg tttaatctgc tgctcttatg 1320
gccaggtttc tttttacttc attatactgg atttgaggac ttcgagtttc ccaataaagt 1380
agtattaatg tgcattatca ttaatggcct tattggaaca gtactctcag agttcctgtg 1440
gttgtggggc tgctttctta cctcatcatt gataggcaca cttgcactaa gccttacaat 1500
acctctgtcc ataatagctg acatgtgtat gcaaaaggtg cagttttctt ggttattttt 1560
tgcaggagct atccctgtat ttttttcatt ttttattgta actctcctat gccattataa 1620
taattgggat cctgtgatgg tgggaatcag aagaatattt gcttttatat gcagaaaaca 1680
tcgaattcag agagttccag aagacagcga acagtgtgag agtctcattt ctatgcacag 1740
tgtttctcag gaggatggag ctagttagct gtctgttgtc tgtagcccag cttgataatg 1800
gaactataca gcgaagagac aatctctggc aagtttttgt agaaaaaatg tttcagtgcc 1860
tagtctgaaa aataacagtt tgagttcttt gaaactctaa aatatatttt tctcatacct 1920
gttttcttca ttttcataat gaagcacttt gctatgtagc tgtgtacata tcactacagt 1980
tataggaagt ttcagtctac agtccatcca aaggaccaac ctgccttaca catctcaagg 2040
aattcagctg ttgaaatcat ttgaactaat caaggaataa atcctaatgt tctgggactt 2100
tattttcaca tgttaaatgc tggaatatat tatgaaaatg ttttcaagaa atcacttaag 2160
tgttcataga ccagtatttc tgacaggtaa aatgctaaaa taagctacct gtaataagtg 2220
tggattatat ttttgggttt tgtagaatat tgcaaattaa ccacacaaaa aatgtttaat 2280
ttatgcaaca agcatgtttg tgcaaatttc atgggacttt aaaaagaata agtatttgag 2340
aaaatatctg gttcacttac actacattta ctgtattatt cttttatagc attaggtgcc 2400
ttgtatttta aatctgtgac aaaccatggc aaatttttaa aggggaagta ttattataaa 2460
atgaagaaat atgtatttct aaaggctata ttgctgtaaa cttaattgat aaagctctgt 2520
ttaatttaga gttttgaaga aatagtctcc cttcaattaa gaaattttca taatggaatg 2580
atttaaattg aagtgacaaa gagtattatt aaaatacaat gtttatacgt gaaaaaaaaa 2640
aaaaaa 2646

<210> 79
<211> 1749
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1483702CB1

<400> 79
tcagcctggc tggcagcagc cttggactcc gcccgtggag ccctgggcct gttgacccac 60
cagcttagga gcacccacca agctctgggt caacgtggag gtaccaggcc accatgctca 120
gtctcaagct gccccaactt cttcaagtcc accaggtccc ccgggtgttc tgggaagatg 180
gcatcatgtc tggctaccgc cgccccacca gctcggcttt ggactgtgtc ctcagctcct 240
tccagatgac caacgagacg gtcaacatct ggactcactt cctgcccacc tggtacttcc 300
tgtggcggct cctggcgctg gcgggcggcc ccggcttccg tgcggagccg taccactggc 360
cgctgctggt cttcctgctg cccgcctgcc tctacccctt cgcgtcgtgc tgcgcgcaca 420
ccttcagctc catgtcgccc cgcatgcgcc acatctgcta cttcctcgac tacggcgcgc 480
tcagcctcta cagtctggtt tcctggagct ggaaagccct gggctcagta aggtcctccg 540
cacaggagcc ttcgcctatc cattcctgtt cgacaacctc ccactctttt atcggctcgg 600

gctgtgctgg ggcaggggcc acggctgtgg gcaggaggcc ctgagcacca gccatggcta 660
ccatctcttc tgcgcgctgc tcactggctt cctcttcgcc tcccacctgc ctgaaaggct 720
ggcaccagga cgctttgatt acatcggcca cagccaccag ttattccaca tctgtgcagt 780
gctgggcacc cacttccagc tggaggcagt gctggctgat atgggatcac gcagagcctg 840
gctggccaca caggaacctg ccctgggcct ggcaggcaca gtggccacac tggtcttggc 900
tgcagctggg aacctactca ttattgctgc tttcacagcc accctgcttc gggcccccag 960
tacatgccct ctgctgcagg gtggcccact ggaggggggt acccaggcca aacaacagtg 1020
aggccccatc cctgaccctg tcctggaggg ggcagaggcc aggccccagt gctgacgagg 1080
agcccagatt tgggcctaat caggtgggga cgcatctcag cctggaacca acaggggctg 1140
aggagagagg gcacaggaga gagggcagag aagaggaggg gtgtctaggg ggactggcag 1200
agtgtgagag ggaccgtgag ggggctcttg atgggagtgg aagaagtgct gagggtctga 1260
gaggggagat gcatgcgtgt ccaggctgaa gatgccccta tattctgtca aaggttggcg 1320
gggggaggtg ttggggtcct ttcatctggc tccgtttctg gtgcttctgg aagtctctgc 1380
tcagcacagg gaagaactaa cacgactaac ctaggcctac cctgaatgct tcttgctaac 1440
caggccgaga ggccacacac ttgccccccc atccccacaa accaggtaat gccagtttgc 1500
cagcagctat ttgcctatag agatgagtct gtcctggtca taactgtgtg ctcaaggtgt 1560
ccaggctttt gggggtgggc ctatctgggt gcattatgga tggtttggtg gattgaggtg 1620
tggggaggag ggtcctaggc tagagggggt atccctagtt agactttggg aagccacctt 1680
caatgttttc tggaacaagg caggtacaaa taaaaaaata aaactttgga aagcacaaaa 1740
aaaaaaaaa 1749

<210> 80
<211> 2339
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1519324CB1

<400> 80
ctcttcatta gttgaaaggg atagtaatac ctacttgcag gttgttgtca tctgagttga 60
gcactggtca cattgaaggt gctgggtaag tggtagctct tgttgcttcc cgttcagcgt 120
cacatctgca gtggagcctg aaaaggctcc acattaggtc acctgtgcac agccatggct 180
ggaatgatga aggggatacg ctggagttgc cctgccatcg cctccatcag ccagacgagg 240
tcctcacagg agaaggacag ctcttcccca ccctgggatc tcaggagggc agccacggag 300
tggggaggcc ccagatgcgc tgtgccaaag ccaggtccga ggccaaagtt ctccctgcca 360
tccttggtgc cgtcctgccc cttcctcctt catgcctggg cctgcaggcc caccccagcc 420
accactgagt ccactcggag tgccctgtgt tcctggagaa ggcattccag ggttgaatct 480
tgtcccagcc tcagcctggg acacctaggt ggagagagtg gtctccgctc tgaattggat 540
ccaggggacc tgggctcatt cttcttggct caccaaccct gcaggcctca tctttcccaa 600
aacccacttt gtcttggtgg gagtgggtcc gcgctgctct gcagcagggg ctggggagtg 660
gacagcatca ggtgggaaag tggagtccac cctcatgttt ctgtaggatt ctcaccgtgg 720
ggctggaaga aaagagcatc gacttgattt ctccaaccac tcatccctct ttttctttct 780
tccaccactc cccaccccag ctgtagttaa tttcagtgcc ttacaaatcc taagctcaga 840
gaaagttcca tttccgttcc agagggaagg gaacctccct aggtccttcc ctggcttgtt 900
ataacgcaaa gcttggttgt ttatgcaact ctatcttaag aactgcccag cctcagctga 960
aaacccgaat ctgagaagga attgcgtcat gtaagggaag ctggaattaa gggagctgag 1020
ccagtcatgg ttgtggcgtg tgagtcagga gacctaggtt tcagcccctc tctactgtca 1080
gcgagctgtg caacgtgggc aagtcattgt cctctgagct gcagtttcct catctgtcac 1140
atcgctacag acaagacctc cctggaaccc ttctgattgt cttagacact gtggttgcaa 1200
aacccacgga aagcctcatt tgtgtggaaa gtcagaggaa aaatgatcca gtggacactt 1260
ggggattatc tgtcattcaa gatccttcct tcaaccccaa ggtcagctcc catctcattt 1320
ccagaaaggc tcatacctgg cttgcaggga agcatctgtc ttgtcattcc aggtgccaga 1380
atcctctcag agtcattgaa gggtgttcac ccatcccacc caaggcttgg cacactgcca 1440
gtgtcttagc agggtcttgt gagggctggg ggcatccagg cactcagaag gcaaaggaac 1500
caccctaccc atttggcctc tggagggggc agaagaaaga aagaaacctc atcctatatt 1560
ttacaaagca tgtgaattct ggcattagct ctcataggag acccatgtgc ttccttgctc 1620
agtgcaaaac tgatgattct acttgctgta gatgaatggt taacacgagc tagttaaaca 1680
gtgccattgt tttgccagtg aagcctccaa ccctaagcca ctgggacggt ggccagagat 1740
gccagcagcc tctgtcgccc ttagtcatat aaccaaaatc cagaccttat ccacaacccg 1800
gggcttggaa aggaaggtat tttggaatca caccctccgg ttatgttgct ccagtaaaat 1860
cttgcctgga aagaggcagt cttcttagca tggtgagctg agttcatggc ttttttttgt 1920

agccagtcct gtccctggcc atccatgtga tggttttgga tggagttaaa cttgatgcca 1980
gtgggcagtg catgtggaaa gtatcagagt aagcctctcc cctccagagc cctgagtttc 2040
ttggctgcat gaaggttttc tttagaatca gaattgtagc cagtttcttt ggccagaagg 2100
atgaatactt ggatattact gaaagggagg ggtggagatg ggtgtggcag tgtatggtgt 2160
gtgattttta ttttcttctt tggtcatggg ggccaaggag aaaggcatga atcttccctg 2220
tcaggctctt acagccacag gcactgtgtc tactgtctgg aagacatgtc cccgtggctg 2280
tggggccgct gcttctgttt aaataaaagt ggcctggaag ctggaaaaaa aaaaaaaaa 2339

<210> 81
<211> 1006
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1630169CB1

<400> 81
ctggcccaag cgtcccctcg tctctctgat ctggcccatc cggcttcgga gggaggcgag 60
ggtgggtgag caaagggatt gggtctgtgg ggtccaggcc cgaacccctg aagacgggct 120
ccgcccccgg cacccgctcg cgccccgccc ccggctggag gagtctctcc tggaccatcc 180
gaacctagcc tgtcccggcc cgcagcctct atggaggctc ctgccgggcc gtagagccct 240
tcgccccctg gggacccacc cgtctataag gtccgtttgg cctgcagcag cctgagtccg 300
taatgctggg cactgttcat gggatcggcc ccctatggag ccctgtgtct ataggggact 360
cctacggtcc ctagggttcg gccccgtcca taatgactcc atatacaggg ccttccatcg 420
ctctataggg ctcagccctc ggcttccaga gcctgtcagc agtggccgta cccttcgccg 480
ggactgccgg gtctccggga cctcttcgat ctacaaggtc atgttatgcc tatagaggtc 540
gcatttgcag ggcctcaccc cgggtagagg gtcctctcca ggtttttacg gcctgtcccc 600
gctctaaggg tcagtgcagg aggcggcgat ggccctgggc aaggttctgg ccatggcact 660
ggttttggcc ttggccgtgc tggggtcgct gtcccctggg gcccgggcgg gggactgcaa 720
ggggcagcgg caggtgctgc gggaggcgcc aggcttcgtg acggatggtg cgggcaacta 780
cagcgtcaat ggcaactgcg agtggctcat cgagggtgag tggggccgcg tgggtcactc 840
actaattcgc tggtagttca ttcatgtgtg cattcatcca ctcaccacat tcccagagct 900
cggttctggt ccaggccttt tgctctaaaa tgcctgacac ccagggttgc agccaacctg 960
gggggatcc actagttcag agcgccgtcc cgcgtggctg cagctt 1006

<210> 82
<211> 1050
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1664253CB1

<400> 82
cgcttccggc acagcggaac tccgggtgcc ggttgaggtt gctggtggac ctgctctggt 60
ggtcttggat gaggccccat gagcgcggcg cccctggtgg gctacagcag cagcggctcc 120
gaggatgagt ccgaggacgg gatgcggacc aggccggggg atgggagcca ccgtcgtggc 180
cagagccccc ttcccaggca gagatttcca gtacctgaca gtgtgctgaa catgttcccg 240
ggcaccgagg aggggcctga agatgacagc acaaaacacg ggggacgggt gcgcaccttc 300
ccccacgagc gaggcaactg ggccacccac gtctatgtac catatgaagc caaggaggag 360
ttcctggatc tgcttgatgt gttgctgccc catgcccaga catatgtccc ccggctggta 420
aggatgaagg tgttccacct cagcctgtcc cagagtgtgg ttctgcgcca ccactggatc 480
ctccccttcg tgcaggctct gaaagcccgt atgacctcct tccacagatt cttctttact 540
gccaaccagg taaagattta caccaatcaa gagaaaacca ggacctttat tgggcttgag 600
gtcacttcag ggcatgccca gttcctggac ctggtttcag aggtggacag agtcatggag 660
gaattcaacc tcaccacttt ctaccaggat ccttctttcc acctcagcct ggcctggtgt 720
gtgggtgatg cacgtctcca gctggagggg cagtgcctgc aggaactaca ggcaatcgtg 780
gatgggtttg aagatgctga ggtgctgctg cgcgtgcaca ctgagcaagt ccgctgcaag 840
tctgggaaca agttcttctc gatgcctttg aagtgagcac cagaggcctt cctcctccag 900
ggccctctgc agaccaggct gagatggagg aacctgctaa aatcgatgga gatgcttcta 960
gcctcccagt aggaggcccc agccatgcct tcaacctggc aggaggtgta gccactcctc 1020

atcctccctg agtgctgata ttctctctct 1050

<210> 83
<211> 1774
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1864715CB1

<400> 83
gggccctggc cgggtggtcc cgggcggatc ccgggaaggc ggaaggcttc ggcagagctg 60
cgccgccgag gctgagcggt cccttctcgc tgcggccgcc caggtgcccg cgcccgtggc 120
gctatggagg cggcgctgct ggggctgtgt aactggagca cgctgggcgt gtgcgccgcg 180
ctgaagctgc cgcagatctc cgctgtgcta gcggcgcgca gcgcgcgggg cctcagcctt 240
ccgagtttac ttctggagct ggcaggattc ctggtgtttc tgcggtacca gtgttactat 300
gggtatccgc cgctgaccta cctggagtac cccatcctca tcgcgcaaga tgtcatcctc 360
ctgctctgta tctttcattt taacgggaac gtgaagcagg ccactcctta catcgctgta 420
ttggtgtctt cttggttcat ccttgccctg cagaagtgga tcatagacct ggccatgaat 480
ctatgtactt tcatcagcgc ggccagtaag tttgcacagc tccagtgtct gtggaagacg 540
agagactcag gaactgtgag tgcgctgact tggagcctct cttcctatac ctgtgcaaca 600
agaataatca caaccttaat gaccaccaat gattttacaa ttcttctacg ttttgtgatc 660
atgctggctt taaatatatg ggtaacagtg acagtacttc gctaccggaa gaccgctata 720
aaggctgaat gatggataca ttattccttc acacagtgga ttttgagtaa ctgaaccaaa 780
ggaaaaagaa gctctttgct aaattaaggt cttttataaa tttagtaaat cagtttataa 840
tctttaaagc caaaggtttt tttagacttg aaagaaagag ccacttaaat tcttgtttaa 900
aaataccaat ttgcctcctc cttcctcact tcgttaggtt atggtagtgc tcagacatct 960
gcagtgttga gaaggccagt cactgttgga agtcatccaa gaagcccatt ttgaggccat 1020
tttgagcctt actcttaagt tctctatgaa gaactacatt gatttgttgg ctttcagaat 1080
cttttaggaa ataaatcctc tccaggacaa aaatgaacat gaatggagtg gcattttgtt 1140
ccaagtcaga ggtgggcacc tataataaat gactagggtt cactttctgg gactgatgtt 1200
taattgtaac acagatacaa cagggtggcc ttgttgtata taatacggta ttatacctgc 1260
atgtgctcta gcaaggatac caaggcaagc atacatgtag ctggcttgag tttgtaccaa 1320
aacagtcctt caactttgca ctgtgcctta agtaattact aacaaaaggt actaggatta 1380
gctgcaatct ctactttcga tgaggaaatc ccagtaagct ttctgattca agtacaatgc 1440
tgccattttt taaagggcca caactataga attaccactg ttggaatttg gtacaaaata 1500
tgttttgtct attgaaaaca tacacggtaa atggtgttgt taggtaggtt ctgtccagtt 1560
cttagggact tttttcacat tatagcattt ttaccctaaa catgatgttg agattattat 1620
atactgtatt ttcttctaaa ttaaccctaa tgtttaaaaa ctcactttcc ccctttaatt 1680
gaaggcattg ttttgttaga tgcagtaatg atgtttacca gagattattg tttcctatgc 1740
aaaataaatt ttcatatttt gaaaaaaaaa aaaa 1774

<210> 84
<211> 2608
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1929395CB1

<400> 84
cgtacggctc gagtctagct tcagactcat gtcatttcca gtctgtcagc gttagacatg 60
gtcctgttcc agtttccaga cattcatttc ttactataga gaagagtact cccctggcgt 120
cttaacctat ggaaaacatg cacggatagg atatatttga ttgcctcctc ttccctttca 180
gtatatgtat tattaatatt attattatta ttattattat tagttcatca gtttgctgtt 240
ctctgcagtg agcagaatca aatgggcaat atttgtcctg ggagacctgt gccgcaccca 300
ggtccccgtg ttaacgtgtg cctgcggttg tggttggcac cctcgggtgg tagctcttct 360
actgtaatga gacaagcctt tcttctgtca ctgcagaatt tagaaggggc cgtgagcagt 420
cttcccaagc atggtccagg agcagctcag agagggtgga gtgaggatgc atgcatccag 480
ggaacaggca tcaagaagcc agggccgtgc ctggagcact gcagagatga ctttggagaa 540
agcaaagcag tgacccagtg accaggcaca ttactcgtga gcgaggtatt cccagtcttc 600

ctgctaagcc taatccaagc cttacccagc tgtgctactg tcatttgcaa agcaaggaaa 660
tactactact ggtaataaaa ctacacatgc aagatgtcag ataagaaata ggtttatatt 720
taccttcctg cgatgtgggg ttggtgcttg aagataaatg tgctgtttct agtttcttaa 780
atagcccaca cccctggatg caaaaaggca gggttatttt tacccatgag catcctgtgc 840
aggcaataga cttccctcac tgtctccgcc tgcagggagg agaatgtacc ccactgcacc 900
tgagctcctg gtccctcagc caagacctca ggggtcccca gccagcctgc tgctggggac 960
ccctgtgctt gctgcagtgt atggagcctc ttgccttccc ctggggaggc acccctgtac 1020
cccagcttcc ttcccctggc cttttctggc cccagtgctc ctcctttaca tagacttgtt 1080
cacacagaaa cgtgcacgcc ccctttttctc cgccacttca ccagtttctg aaatccaacc 1140
tcccagactt cacaggaaga tagatattct tgagataatg aaaagtgata tcttcgcata 1200
cgaaaggaaa aaaggttgag gtatatatga tttttaactg tattaggggt gtatgaacca 1260
gtttaaaaac gaggttttat ttactgtaga gatgaatgca aatcagaacc aatgatccct 1320
tggcctactt agttaaaacc agttcataca tcccttaggg tttttattat tattattatt 1380
attattacag ttgttattgt tgtttttgtt gttattatta tttggggttt cttgtgtttt 1440
ttctttgcga ctctccacac taaacttgca atattgtggg gagaagctgt gactaaactc 1500
tacgctgcgg tgagatgtag cagtaatcag ctcccagcga cgtgtgtagc tggggctgcc 1560
gctcgcaata atcactattg atttaaagct ttacttagcc ttgatctgta ccctcgtagt 1620
caataaggtc ttgccacatt ttattagtga ggtggagaaa cgtattattt gtttgttgtt 1680
tttgcccttc ccccacccca caatattaaa ctgtgaaatt tgtgatttgt ttaaactctg 1740
ggtgaatcat agcttagttt gcatgtccag ctaatttgtt tctatacatt ttgtttgatt 1800
ctctttctcc ttctctcagg gcttttacaa aaaaatatat atatatggat cttctgaaaa 1860
gtttttttgag gtgcaagttt tttctcttgt tttttttttt ttttttttct cattgattaa 1920
tggacatgat gctgagattc aatcactaca tgaaacacct ggctgtgaaa acaaaacaac 1980
ccagagggct gtgttccaag cagcgctggg gaagctacgt aacagtcgga tgccagtttt 2040
ggaagattca ccatgcgttc tgaccctctg ttcgtctctt tcctctcctc tttcttcaag 2100
aaggaaattg atcctagtga tttcagccca tgcattaaac aggaaacaat aataaatttg 2160
tagaattcat atttttctaa agggaactta aaaactgctg ctacatgtta tgtacaaaac 2220
tggtttatgc cacatgaaca gagaatcaca agtttggttt tggtactttt tgttcctctt 2280
tgtattcagt tgtatagaac ttccaaattc agaatgagaa gaaagctgtt ctgtatcaaa 2340
ccatttaagc aataaatgtt atattttaaa agctgcagaa tactggtaga tgctgccatt 2400
gttgatatgg ttcctggcaa acccatgtga tttgagagct tctcagacta tccacctttg 2460
ggtcgctttg ctgttcgtga tatgagacag acagttgcgg tgggtgtcat caaagcagtg 2520
gacaagaagg ctgctggagc tggcaaggtc accaagtctg cccagaaagc tcagaaggct 2580
aaatgaatat tatccctagt acctgccc 2608

<210> 85
<211> 1336
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1987737CB1

<400> 85
cccacgcgtc cggtgacttt caaatcccct ttcgcagtga aagaaacagc aaacacttaa 60
gattcagcat ctgttctcca gttgcactga ggaatgcact gtctcgcagc accagctctg 120
cagagccctt gccccagact ctttgcggtt ttatttatat gtatttccat atttcattcc 180
tgtgtgtcac tgctgcattg gtgtggcagc aagtgaccaa atgctacagg tcttactatg 240
gacaccaggt caggtgcaac cacacaaaac aaagccagtt ccatgagctg cctatgatat 300
gcattgcgga agtaacattt tacccagggt gtgccattgc agtgatataa atatattttt 360
ttcttagact aaatatgagc tgactatctc ttttgatgtg tgtacatagg tgtgagtgtg 420
tctgtatgcg tgcctgtctg tgtgcgggtg tgtgtatgtg cgtagcctca tgcttaggac 480
tacccatgaa tgttgtggaa tgctacacct ggagagttct ggttttccac cagtttcaag 540
atgaagaact acatgataca gtggacctgg agaccatccc cttggaaaga caacccagag 600
atgttcagca tcctgtatct acacgcatcc tgtatctaca cgtgtatttt gtagctgtca 660
cactaacctt aataagaatt ctacagcttt ggacagaggc attttcacct taatggtgaa 720
gtaatttaaa atataaatcc attcaggtga caacccatca tcaaaattac aaattttctg 780
attgaactca tctgaatcat cagttccttg atggagagag agaaggagat ggaatgtgtc 840
tggtaacccc aaatggagta caagtagcct ttgttttcct gcataaatgg acttgttgaa 900
tgcgaacgaa tatatgcaat tcatatactt ttggagatga acgtagatat gtgtgtcagc 960
tttgagatgg tgtgtcctgg attaatactt tgtctcccaa tatcacagaa aaatacatgc 1020
cagtgactct tgaggttaag gtagttggga tgaaatggcc tcaggcaatt tcacattccc 1080

taattacctg gaaagttcta cagtaattaa tatgcagcta actcctgttg ccctcacaag 1140
agcatcagcc ttctagaatc ggagctccgg agtgtgaaga ttcagtattg atatgatatg 1200
tataccaaac tccagccaac ttactgccat ttttcataat ctgagtggct gccttgctta 1260
tcctaagctg tggttgcaga aaccgtggcc atttatataa gctataacat caaatcaggg 1320
aaaaaaaaaa aaaaaa 1336

<210> 86
<211> 3062
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2122866CB1

<400> 86
gtcccgccag ctgtgaggag ccccgagcac ggctgtcgcg gcctgaggca gagcgcctct 60
gctctggccc tgtccctggc tccctcatta ccgtcgccag cggagcgaac ccgagagcgt 120
ggagcgggct gggccagctg ttgagtggaa tgtcatggcc agctcctcgg acagtgaaga 180
tgacagtttc atggctgtgg accaggaaga aactgtgctg gaagggacaa tggatcaaga 240
tgaggagccc cacccagtat tggaggctga ggagactaga cataataggt ccatgtcgga 300
gctgccagaa gaggttttgg agtatatcct gtcctttctc tcaccgtatc aggaacacaa 360
aactgcggcc cttgtctgca aacagtggta tcgacttatc aaaggtgtag cccatcagtg 420
ttatcatggt ttcatgaagg ctgtccagga aggaaacatt cagtgggaga gccgtaccta 480
tccttatcct ggaacccccaa tcactcagcg cttctcgcac agtgcatgct attatgatgc 540
taatcagtct atgtatgtgt ttggaggctg tacccagagc agctgcaatg ctgctttcaa 600
tgacctctgg agacttgacc taaacagcaa agagtggatc cgacctttgg cttcagggtc 660
ctatccttcc cccaaagctg gagcaactct ggtcgtgtac aaggacttgc tagtgctgtt 720
tggtggctgg acgcggccaa gcccttatcc cctacaccag ccagagagat tctttgatga 780
aatacacact tactcaccct ctaaaaattg gtggaactgc attgtgacaa cccatgggcc 840
acctcccatg gctggccact cctcctgtgt gatagatgat aaaatgattg tctttggtgg 900
ctctttagga tcccggcaaa tgagcaatga tgtctgggtc cttgaccttg agcagtgggc 960
gtggtccaag ccgaacatct ctggccccag tcctcatcct cgaggtggcc aatctcagat 1020
tgtcatagat gatgcaacta tcttaatcct cggagggtgt ggcggtccca atgctctatt 1080
caaggatgct tggttgttgc acatgcattc tggtccttgg gcctggcagc cactcaaggt 1140
agaaaatgaa gagcatgggg ccccagaact gtggtgccat ccagcttgcc gggtgggaca 1200
gtgtgtggtg gtcttcagcc aggctcctag tgggagagcc ccactcagcc ccagtttgaa 1260
ctctcgccca tcacctatca gtgccactcc tccagctctc gttcctgaaa cccgagagta 1320
ccgctctcag tctccagtaa gaagcatgga tgaagctcct tgtgttaacg gccgctgggg 1380
aacactgaga cccagggctc aaaggcagac tccttcaggt tcccgggaag ggagcctttc 1440
cccagccaga ggagacggct ctcctatcct caatggtggg agtttgtctc caggaacggc 1500
agctgtgggt ggctcttctt tggacagtcc tgtacaggcc atatctccaa gtactccatc 1560
tgctcctgaa ggatacgacc tgaaaatagg actttctttg gccccccgac gaggatcact 1620
accagatcag aaagatctga gattaggatc catagatctg aattgggatc tgaaacccgc 1680
ttccagtagt aatcccatgg atggcatgga caataggaca gttgggggaa gtatgagaca 1740
ccctcctgaa cagacaaatg gtgtgcatac cccacctcac gtggccagtg cccttgcagg 1800
ggccgtctcc ccaggtgccc tgcgtcggag tctggaagcc atcaaagcga tgtcctccaa 1860
aggcccctcg gcctctgcag cactaagtcc tcctcttggg tcttctccag gctctcctgg 1920
gagccagagt ttgagcagtg gagaaacagt gcccatccct cgcccagggc ctgcccaagg 1980
agatggacat tccttacctc ccattgctcg ccgcctgggc caccaccctc cacagtccct 2040
aaatgttggc aaacccctat accagagtat gaactgcaag cccatgcaga tgtacgtgct 2100
ggacattaaa gacaccaagg agaaggggcg ggtcaaatgg aaagtattta atagcagttc 2160
tgtggttgga cctcctgaaa ccagcctgca taccgtggta caaggcaggg gtgaactcat 2220
catatttgga ggactcatgg acaagaaaca gaatgtgaag tactatccaa aaacaaacgc 2280
cttgtacttt gtacgagcaa agagataatg tgttctaaac ccctttcctt ttctgtggct 2340
tttaatttgg aattttccag tgtgtaagca tttggactga gaattgggaa aacaaaatta 2400
ctcccagaag ccaaaactct ttaattccca accgaagtca ctccaggctg ggatcaaatc 2460
tccattaaga aaaaaaatta tatataaata tatatatata tattatatag ccaactctgt 2520
tgacaaaaaa agggagagat ttccatcctg gttcagataa agttgttgct gtgttttaac 2580
aggggctggg ctgccttttt ctaccttgct ggtaactaga ccaagaagtt agagaataga 2640
ctaacatcag taacttccca aaagaaactg aagagccccc tgtaaatctt tatgtggcct 2700
tcttggagtt aaaaaatgaa agggcatatg taagttgcaa aggtggaggg ttttagactc 2760
tcatgcttca ggtgctgtcg gggtaaaagt aactgttttt ccccttctct taaaaccaca 2820

gaggacctgt gacagctctg cagaaatgcc agtgcctggc cccctcttgc cttttatggc 2880
tgaggaaagt tacccaacaa aggattttat tccacatttg tgtgccgggt cattgtgaaa 2940
taatgtttat gcagccaaca tctgaccgcc tagtagtgtc cattggtctt tggagtgctt 3000
cttgtgtgtc tcagaaaaca ttttgtgtct gattgtggaa atttctgaca atcaatttct 3060
tc 3062

<210> 87
<211> 2543
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2123981CB1

<400> 87
gcacctgcga gcctggaggt ggcggtcgct gcgaacccgg cggcgcggcc aaacctgtgc 60
tggctgcctg ggccttctgg gcctcctggg ccagaagctg gcgctggggg cggcggcggg 120
gtgggaaagc ggcgccgccg tggggcctcc tggacgcgtc ggcggaggtg gcttcgtccc 180
gcggagtcca ggcttcaggc tttcaccagt tctcaggatg cccataggga tgggtgaagc 240
ctgcctggcc tgtggtgctt cccagtggcc gtcatctcat tagggcccca cagtggcatt 300
aggatgcacc tctcggcggt gttcaacgcc ctcctggtgt cggtgctggc agcggtcctg 360
tggaagcatg tgcggctgcg tgagcatgca gccacactgg aggaggagct ggccctcagc 420
cgacaggcca cagagccagc cccagcactg aggatcgact acccgaaggc actgcagatc 480
ctgatggagg gcggcacaca catggtgtgc acgggccgca cgcacacaga ccgcatctgc 540
cgcttcaagt ggctctgcta ctccaacgag gctgaggagt tcatcttctt ccatggcaac 600
acctctgtca tgctgcccaa cctgggctcc cggcgcttcc agccagccct gctcgaccta 660
tccaccgtgg aggaccacaa cactcagtac ttcaacttcg tggagctgcc tgctgctgcc 720
ctgcgcttca tgcccaagcc ggtgttcgtg ccagacgtgg ccctcatcgc caaccgcttc 780
aaccccgaca acctcatgca cgtctttcat gacgacctgc tgccactctt ctacaccctg 840
cggcagtttc ccggcctggc ccacgaggca cggctcttct tcatggaggg ctggggcgag 900
ggtgcacact tcgacctcta caagctgctc agccccaagc agcctctcct gcgggcacag 960
ctgaagaccc tgggccggct gctgtgcttc tcccatgctt ttgtgggcct ctccaagatc 1020
actacctggt accagtatgg ctttgtgcag ccccagggcc cgaaggccaa catcctcgtc 1080
tcaggcaatg agatccggca gtttgcacgg ttcatgacag aaaagctgaa cgtgagccac 1140
acaggagtcc ccctaggcga ggagtacatt ctggtcttta gccgaaccca gaacagactc 1200
attctgaatg aggcagagct gctgctggca ctggcccagg agttccagat gaagacagtg 1260
acagtgtccc tggaggacca cacctttgct gatgtcgtgc ggctggtcag caatgcctcc 1320
atgctggtca gcatgcatgg ggcccagctg gtcaccaccc tcttcctgcc ccgtggggca 1380
actgtggtag agctcttccc atatgctgtc aatcccgacc actacactcc ctataagacg 1440
ctggccatgc tgcctggcat ggacctccag tatgtagcct ggcggaacat gatgccagag 1500
aacacagtca cacaccctga gcggccctgg gatcaggggg gcatcaccca tctggaccgg 1560
gctgagcaag cccgtatcct gcaaagccgt gaggtcccac ggcatctctg ttgccggaac 1620
cccgagtggc tcttccgaat ctaccaggac accaaggtgg acatcccatc cctcattcaa 1680
accatacggc gcgtggtgaa gggccggcca ggaccacgga agcagaagtg gacagtcggc 1740
ctatatccag gcaaggtgcg ggaggcacgg tgccaggcgt cagtgcatgg cgcctccgag 1800
gcccgcctca ctgtctcctg gcagatccca tggaacctta aatacctgaa ggtgagggag 1860
gtgaagtacg aggtgtggct gcaggagcag ggggagaaca cctacgtgcc ttacatcctg 1920
gctctgcaga accacacctt cactgagaac atcaagccct tcaccaccta cctggtgtgg 1980
gtccgctgca tcttcaacaa gatcctcctg ggacccttg cagatgtgct ggtgtgcaac 2040
acgtagcgag caggccacag cctggcctcg ggaaggtggc tcctgcagtt cagcgtccct 2100
gggcccatta atcccactgt ggagacttct gggaactatt tattgagcag gcctgtgcct 2160
ccacatcatc ttgttgtctc tggggtgtgg tgtcacagca ctcctctttg ccctagagat 2220
aagggacctg acttcccctt ctcccatcct gaacatttgt acccctggag aagttcctta 2280
gcagggagga ggaagaggag aggaggaagc aaagaatcac aaggaacctc tggctaggtg 2340
atcctgatgt ttcctactga gtttttctgg tatccagatt tctggaaacc gagtaatcat 2400
gtactgtttg attgggtggt tcatctgctt ccatcccagt gaaatttacc tgtagcccag 2460
tgaagggtgt gtttggaaca ttcaaaatat cacttcaaca tcttatcaaa ataaaaatta 2520
ttaatgagaa aaaaaaaaaa aaa 2543

<210> 88
<211> 873
<212> DNA

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2200177CB1

<400> 88
tacatgcgtt aatcactttt gcacacctgg ataatatctg cttgctttat gatctgtttg 60
aagaaataaa cgttacagta acaagaaagt ctgggaaggt ggtacctaaa aagaaacaaa 120
acagaaaaca aaaaacaaaa acctcagaaa cttgtttggt aagtaacatc cttttatata 180
acattgataa tctctattct tgtctgtctc tgtgtcacaa aagaagataa ttgtagagct 240
tgaatgatat atcaaagatt atctagtaca tttttcacag gtacagaaac agatgctcag 300
tgaggataaa ttatttcaaa taatacacag tctgttaata acacagctag catcagaaac 360
taagatctct gcaaccatct gtctgccttt actcttccat tgcttgtttc tgttagtttt 420
gagctttcct atcacactct gtataaggca ctctggcccc tatcatattt atcctctgtt 480
acaggtgtca aatctcattt tcctccagac ccatttcctc agctacattg ctggcattat 540
gcaaaagctc ctttccaatg ttgtgcactc tcaaaaaatt catccagaaa tactaaggtt 600
tgggaaggtc tgtgcccaaa gcactataag caaaaagttc aaagaagaaa aatacaaaac 660
tcctcataca ataagtctaa tatcacaaat tcatgaaacg gcaactataa agagcaaagt 720
attcagaaaa ttaagtacat atttcagcat tgtcttaaag ctaaaagaaa ttaaaatagc 780
tggttttaaa tatctatgga gttcaaacta gattagtgtt gaactcttca ccttatcagt 840
atatatataa taaaactttg gttatgaatt tat 873

<210> 89
<211> 2134
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2319255CB1

<220>
<221> unsure
<222> 713
<223> a, t, c, g, or other

<400> 89
ctctgcaggc acccaccgtt tctgttttct tctgttatcc ttccagagat tttctatgcc 60
taaacaccat ttgaagccat gaatgtatag tgatgagggg ttttttacag ttgcacatga 120
ctgggaatag aaatgcaatt agatattgag aatattgatg agccaatagg agctgatggg 180
ttagaaaaat aaggaggtga agagaatcca ggattttaag gtttaaatga agttagagac 240
tagatttgga gcgagtaaga gaagatacag aattctattt aaaagattca gaactgggaa 300
aattagaact agcattagtt ttaactgaaa tcatataatg catgtgttca taccaagtag 360
aacaaaaatg gccttttttt ctctcatatt taaactccac attatctact gtaatgttcc 420
acaaagagca catatttagt aaatattgcc tttagtgtcc aattatgttt gagcttgctg 480
tagagtttag atttctaatt gctaatcatg tttcattttt aaaataatct ataatgttat 540
ttgcaagaat ttcaatccat agcctcctag aaaaaaaatc aataccatag gcagaaaaat 600
aatatttcag gtaatagagc atatcactgc atacaaatcc attcttactt caatcaaaaa 660
agaatatgat gcctttattg agacaataaa gaaagaccga agaactacat ttngtcttca 720
tggaaaactt aaaggtttgg cagcagagcc tacagctttg gtatattaca ggaaaagaac 780
aatccaactt gaagcaaagc atgactcttc aagaatccat gaatctagat gttctcacta 840
aatacatgaa acatcttgaa gataaatatg cagaaattaa acaagctatg ttgataaaat 900
atgtgccagc tcagaggaag gctgatttag atgaagaaat gattgtgtta ttaaaacggc 960
gagatgtagc tgaaaatctg aacaaaaaat tgcagttttg tcatcaaaga ctgcagataa 1020
tttcacaggc acttagttca tgggtaaaat ctgatatgag cagcccattt caagactttg 1080
tggagcaaat tcagaaaacc aaatatttac aaggtttaat gagttaatct cacttggtga 1140
atatgaaaag gcagcttgtt atgcagcaaa cagtcctaga agaattcttc gaaacattgg 1200
tacaatgaat acatttaagg ctgttggaaa aattagagga aagcctcttc cattactctt 1260
atttttgag gccctcttta tcacaagtca tgcttttcca tgtcctgttg atgcagctct 1320
aaccctggaa ggaatcaaat gtggattatc agaaaaacgg ttagatttag ttaccaattg 1380
ggttacacag gaaagactga cattttctga ggaggctggg gatgtgattt gtgattatgg 1440
ggagcaggat acttataaca aggccaagtg cctggcttta gctcagatta tctacagtga 1500

atgtggcctg cacaagaaag ctattctttg cttgtgtaaa cagggtcaga ctcatagggt 1560
catggagtac atacagcagt tgaaggactt tactaccgat gacctgttgc agctattaat 1620
gtcatgtccc caagttgaat taattcagtg tctcactaaa gagttgaatg agaaacaacc 1680
atctttatct tttggtcttg ctatacttca tctgttctct gcagacatga aaaaagttgg 1740
cattaagcta cttcaagaaa tcaataaagg tgggatagat gcagtagaaa gtcttatgat 1800
aaatgattcc ttttgctcca tagaaaagtg gcaagaagtg gcaaatatat gttcacagaa 1860
tggctttgac aaattatcta atgacatcac gtctattctt cgatctcagg ctgcagttac 1920
agaaatttct gaagaggatg acgcagtcaa cctaatggaa catgtgtttt ggtagttcta 1980
tatcttaacc agctgaggga gcttgtacaa caccttatgt atgctggttg ggaaaaactg 2040
attttaacgt aactaactta taaataaatc tagagtaaga agatctcaga aataaatcac 2100
acttttgtta tgatgacaaa aaaaaaaaaa aaaa 2134

<210> 90
<211> 1886
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2792452CB1

<400> 90
tggctgtccc caccagctcc aaagtctcca cctctcaaac accatgtctt ctttggtgtc 60
tgggacacag gctgttcttg gccaggccac agcctcatca tgcatggctc tgcccggtgg 120
ctcacgggct ctgtccacaa caccgggga caggtgctct tatctctccc aggagtcagg 180
atgagacgtg catgcatgta ttgccagctt ctcttgggat cctgggtgct tcctcagccc 240
ttggccgtcc cttgggctgg gctccccacc tctctggcac ccaagactgg tggaaaagga 300
aagccctctc gttctgagtt gagatctcag aacccttggc ctttctgcag ggcctccaaa 360
ggcgtctggt ggaccctgca cctgctctgc agcgcccccc ccaggccttc ttgcacggga 420
cagtgacagg gagttctgtc tcagagcgag tccagcaggg cctccacaga gtgcaaggtt 480
gcctgccctg cctgtggaag agcaaacctt gtcccctgtt tcccagaccc cctactcagt 540
ggcctctgaa tgtccccaat gtgggccccc ccctggctgc ccctcttgct ctccaagtcc 600
gagcccaccc agagcccaag ccctcggagg cccctgccac ctggaaagat gacactgggg 660
cagggttctc ttctgatgtc tgtgttttgt cttgtgggcc ttggggtgcc cctgcccctg 720
ataagaaggg gtttcagagc agaaataaag cctcagacag gtgagcctct gtggcacatg 780
gcccctcgag ccagccatgc atctgggttc agcccctgcc aggacacttg accaggcccc 840
ttaccttctc cgagctggga tcatgctggg aagactgaat gggacgcagc cggtgaaaac 900
gcctggaaga gagtaaacgc cattttctct aagcatattt atttctttaa caacagtgac 960
gaccatttgt gagccatctc tctgtctcag gcacggtgct acatgccaac gaaacctgct 1020
cccattgaac cctggccagc cagtgaagaa agggttgggc ctgggaggtg ccactttaca 1080
gacaggggca ccaaggggca gggtggcagg aggcccaccg gacgttcccc atgaagtagc 1140
agtcccagca tccacaccca gcaggcacca cgctggcccg cagcctccct gccagcacgc 1200
ctggcttccc ggcctcggaa cttgatctgc tccctcttcc ggacactggg gctcctgcca 1260
agtcctgggc tgggcagcaa ctgctgaaca ttctaagaaa tccctcccag ggttttctca 1320
ggagcccggg tggggcagga agtccccagg ggctgagggg accgtggcgg caggtggcac 1380
ccagagcagc actctcctgg ggcccaggct gttgggccag aggcaggact gtgaggccta 1440
gtgtagggcc tcctgccagt ggccggcacc tacttgtggg gctgggggtt cccccagcag 1500
gttgggctcc ccacctgaca cactcacaga ccttgtgcct tggagagcca gtgttcccgg 1560
ggccacatag ctatgccgcc caggggctgg gcctgtccca gctctggtcc cccggcccca 1620
ggtcctggac gctggctccg cgcagcagca ggcggcctcc ggaggacacg atgtgactgg 1680
ctgccgctac gtcgcactca gatgagtctg cgccggatcg acctgctgcc gagtcctgcc 1740
ggacaggcac aggcagggag tgaaaattat ctaccccttt ttatttctta ataactgaat 1800
gaaaataaac attggtggtt tgacaaataa ctacatattt tcaaacccag ccagtccagg 1860
ggatgcagtt tccaggtgcg ttatgc 1886

<210> 91
<211> 928
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2853088CB1

<400> 91
ctacaagaag ctattttgcc actaagagag gagagcctgc ctgagggtga agccaaatgc 60
atagggaagg cagaaccaga gatggtgagc cacaggttcc tgatatcttc tggcccctgt 120
gttcacctgt ctgaatctgt ccgctctgga cttttcgatt acttgggcca tagcttcctt 180
tactgtttag gccagctgga gccgggtcaa tcagacactg gatgggtgga ggatctcact 240
gtgctcatgc tttcaactac cctccgtact ttaattgttc tcgatttctc taactggcct 300
tgatgtctct tcttagctcc agactatatg cccacagacg tcccaaaagc acgtctggaa 360
ttgacctcat tattattatt gttattattt ttgagatgga gtctcgctct tttgcccagg 420
ctggattgca gtggcgcagt cttggctcac tgcaacttcc gactctgggg ttcaagtgat 480
tcttctgcct cagcctcctc tcaagtagct gggagtacag gcgcctgcca ccaagcccgg 540
gcaaaagagc gagactccat ctcaaaaata ataacaataa taataatgag gtcaattcca 600
gacgtgcttt tgggacgtct gtgggcatat agtctggagc taagaagaga catcaaggcc 660
agttagagaa atcgagaaca attaaagtac ggagggtagt tgaaagcatg agcacagtga 720
gatcctccac ccatccagtg tctgattgac ccggctccag ctggcctaaa cagtaaagga 780
agctatggcc caagtaatcg aaaagtccag agcggacaga ttcagacagg tgaacacagg 840
ggccagaaga tatcaggaac ctgtggctca ccatctctgg ttctgccttc cctatgcatt 900
tggcttcacc ctcaggcagg ctctcctc 928

<210> 92
<211> 962
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2949004CB1

<400> 92
ttggacctag caatgccatg aaattaccgc gagaggaagc aagcgcaagt ttcgtgagga 60
gggcagatct cacgagagag gatttggcgc cctcctctgt ggattctggc caggccgggt 120
tcggcggttg ctgtgagagc gggcttccca acaccatgcc gtccgccttc tctgtcagct 180
ctttccccgt cagcatccca gccgtgctca cgcagacgga ctggactgag ccctggctca 240
tggggctggc caccttccac gcgctctgcg tgctcctcac ctgcttgtcc tcccgaagct 300
acagactaca gatcgggcac tttctgtgtc tagtcatctt agtctactgt gctgaataca 360
tcaatgaggc ggctgcgatg aactggagat tattttcgaa ataccagtat ttcgactcca 420
gggggatgtt catttctata gtattttcag ccccactgct ggtgaatgcc atgatcattg 480
tggttatgtg ggtatggaag actttgaatg tgatgactga cctgaagaat gcacaagaga 540
gaagaaagga aaagaaaagg agaaggaaag aagactgagg ggcagcagct gcttggagtt 600
tgcgtccttc ccgtccaccc agtgcagctc ccagtgctgc agtgtgcgtg gcgtgggcat 660
ccttccagct gactcatggt ttgaaaaacc gttgttttat ttaaatatcc acagtggtag 720
ggcacacact gaagttggct tttcagccag cactgaatgt atccatcagg acatgcgtct 780
tcaggtgcct gatctttgta gtcaggctgt gggaacggtc tctgcagagc ttcataactg 840
ggaatttgat ttgaagaagt ccatgtcata tgtgtaacta gtactaatta taaatataaa 900
atacacaata taaaatatga aactcaataa taaacagtgc cacctgtaaa aaaaaaaaaa 960
aa 962

<210> 93
<211> 2644
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3011670CB1

<400> 93
ccctctcttt cgctgtttga gagtctctcg gctcaaggac cgggaggtaa gaggtttggg 60
actgccccgg caactccagg gtgtctggtc cacgacctat cctaggcgcc atgggtgtga 120
taggtataca gctggttgtt accatggtga tggccagtgt catgcagaag attatacctc 180
actattctct tgctcgatgg ctactctgta atggcagttt gaggtggtat cagcatccta 240
cagaagaaga attaagaatt cttgcaggga aacaacaaaa agggaaaacc aaaaaagata 300
ggaaatataa tggtcacatt gaaagtaagc cattaaccat tccaaaggat attgaccttc 360
atctagaaac aaagtcagtt acagaagtgg atactttagc attgcattac tttccagaat 420

accagtggct ggtggatttc acagtggctg ctacagttgt gtatctagta actgaagtct 480
actacaattt tatgaagcct acacaggaaa tgaatatcag cttagtctgg tgcctacttg 540
ttttgtcttt tgcaatcaaa gttctatttt cattaactac acactatttt aaagtagaag 600
atggtggtga aagatctgtt tgtgtcacct ttggattttt tttctttgtc aaagcaatgg 660
cagtgttgat tgtaacagaa aattatctgg aatttggact tgaaacaggg tttacaaatt 720
tttcagacag tgcgatgcag tttcttgaaa agcaaggttt agaatctcag agtcctgttt 780
caaaacttac tttcaaattt ttcctggcta ttttctgttc attcattggg gctttttga 840
catttcctgg attacgactg gctcaaatgc atctggatgc cctgaatttg gcaacagaaa 900
aaattacaca aactttactt catatcaact tcttggcacc tttatttatg gttttgctct 960
gggtaaaacc aatcaccaaa gactacatta tgaacccacc actgggcaaa gaaagtatcc 1020
ctttaatgac agaagccaca ttcgatactc tgcgactctg gttaataatc ctgctgtgtg 1080
ctttgcggtt ggccatgatg cgtagtcacc tgcaagctta tttaaattta gcccaaaaat 1140
gtgtggatca gatgaagaaa gaagcggggc gaataagcac ggttgagcta cagaaaatgg 1200
tggctcgagt cttttattat ctttgtgtca ttgcactgca gtatgtggcg cctctggtaa 1260
tgctgcttca cacaactctg cttttgaaaa cactaggtaa tcattcctgg ggtatttatc 1320
cagaatctat ctctacctta ccagtggata atagtctact gtccaattct gtttactctg 1380
aattaccatc agctgaaggg aaaatgaagg taactgttac acaaataaca gtggcactga 1440
gcagcttaaa aaatattttt actcctcttc tttttcgagg acttctgtct tttctgacct 1500
ggtggattgc tgcttgcctc ttttctacaa gccttttttgg gcttttctat caccagtatc 1560
tgactgtggc atgaatctca gttaacaaaa aagcatatcc aaatcaccct ttaaattaaa 1620
atatctgtgc ccttaaaggg ctgatgaaaa ccagaagaaa gcaaatacaa tgggaaaaaa 1680
aaaacatatc agaatgtctt gtattaaatg tttcctctgt attctcaggg tgaattaatg 1740
tagtaatatt taaaattaca aaatagattg ttaactgtta cactgtggca ttggaatttt 1800
aactctttgt atttactggt atgagagggc tatctacaag ggtaatattt ctgattaccc 1860
tggtttacag aaacctccag cagtctttga aacatctcac atgactctag ttattgattg 1920
cttttaatgg ttttatggta ctgttgatag tcatagtggc tgcctataga acaatcttca 1980
aactgagcca tgctttaggg gagggaaagg ggctaaagtc tcttctgttg gtaatttatt 2040
agttactctt gaaacagtaa aatccaacag aaaggaagag atagctactg tatattacag 2100
taaagaaagc tgcatagtta ttttaaattt aatggagatg aatatggtta aaatatataa 2160
ctactgctgc ttgagaatag caagagtatt gttttaaaac atattccacc caacttgaga 2220
gttcttttaa aatgattggc catatgaaca tttgtaatct tgccattagg tttggacctg 2280
ccatattttg ttttattctg tgatcctaac tagttccttt taataggcta aaatatttat 2340
caatactgat cagactttaa agaaattact ttgtaaacct gctgactacc tgtatgtatt 2400
gtatatatat tatatattaa atatataata tattgagatt ataaaagatg aaaatattga 2460
atccttataa tattttaagt tgcagaatgt atgttaaaaa gtgacttgaa tgagatgtat 2520
ttgtatctag aaattttatt tctttttgga atgagattaa aatacatttt gaaagttcaa 2580
caaacagaaa agaaaagaca gtagcagaag tcggtaagac ggagtgtgcg cgagagtagc 2640
ccct 2644

<210> 94
<211> 1875
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3242083CB1

<400> 94
cgtcgcccaa ggtcgcgggc cgcttgggga gtcagcagcg cgccaggccc cttcgggccc 60
cacacgcatt aggtgccttc ttgatgggta cggagtgaac gcgggcggcg gcgggaccga 120
ggcagcgccc agtttgtaac cgccgcgccg cccgtgcccg cgcgcgccac accccagcgc 180
gcttccggcc gggccacgtg accgcgcgtg cacgtgttcc ggcctctccg cttcgccgct 240
ccgaacctcc tcctggtcgt cccggcattc gtccacgcgg agccggcttg ggcggggccc 300
gggaggcggc ggccggagaa gccgcggaga cgcgagcgcc gagcgtcgcg agggagcagg 360
cccgggcagg caagcggcgg cctccgccat gaaccccagg ggcctgttcc aggacttcaa 420
ccccagtaag tttctcatct acacctgcct gctgctcttc tcggtgctgc tgcccctccg 480
cctggacggc atcatccaat ggagctactg ggccgtcttt gcccccatat ggctgtggaa 540
gcttctagtc gtcgcaggcg cctccgtggg cgcgggcgtt tgggcccgca accctcgcta 600
ccgcaccgag ggagaggcct gtgtggagtt caaagccatg ctgatcgctg tgggcatcca 660
cctgctgctg ctcatgttcg aagtcctggt ctgcgacagg gtggagaggg gcacccactt 720
ctggctgctg gtcttcatgc ctctcttctt cgtgtccccc gtgtccgtgg ctgcctgcgt 780
ctggggcttt cgacacgata ggtcgctgga gctggagatc ctgtgctcgg tcaacatcct 840

gcagttcatc ttcatcgccc taaagctgga caggattatt cactggccgt ggctggtggt 900
gtttgtgccc ctgtggatcc tcatgtcgtt cctttgcctg gtcgtcctct attacatcgt 960
ctggtccctc ctgttcctgc ggtccctgga tgtggttgcc gagcagcgga gaacacacgt 1020
gaccatggct atcagttgga taacgattgt cgtgcctctg ctcacttttg aggtcctgct 1080
ggttcacaga ttggatggcc acaatacatt ctcctacgtc tccatatttg tccccctttg 1140
gctttcctta ctaactttaa tggccacaac atttaggcga aaggggggca atcattggtg 1200
gtttggcatt cgcagagact tctgtcagtt tctgcttgaa attttcccat ttttaagaga 1260
atatgggaac atttcatatg atctccatca cgaagatagt gaagatgctg aagaaacatc 1320
agttccagaa gctccgaaaa ttgctccaat atttggaaag aaggccagag tagttataac 1380
ccagagccct gggaaatacg ttccccccccc tcccaagtta aatattgata tgccagatta 1440
aactcctaga gaggacccag gcacacacag actccacttg gccttcgcct cttgttcatt 1500
catcccaaac ctggaaatgg aaacaggctt caaacactcg tctcacgccg tgtttgagat 1560
caccgcctca tcagtatgca tcatagatgg aggtggtttc agtatgtggg tgtgtgtggt 1620
gtgtacctgg gtaagagact tgctttccag gttcgcactt tcaggtgtag ctgggggcag 1680
taagtcgaat tgttttagta ggtcctcaaa aggaataacc acacagctgt ttgtttaaat 1740
gctactgtac ctatcaaaac tattgtttaa aaagtatttt tatacactgc taatctaaaa 1800
ttgtatttca gattgtgcct gtcataacaa tagcaaatgt aaaaagttct ctttcccacc 1860
aaaaaaaaaa aaaaa 1875

<210> 95
<211> 2378
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3363391CB1

<400> 95
aaagtgcagc agatgttgcg ttgctagttc aattcactgg gcggccgtaa tatggcgttg 60
gacgcgtgcg gtaaacgacc gtcgtagtag agtcagagca catgacgggc gaggcaagga 120
atgtgttctt ggtctctcta ccatgacctg ggaccgatga tctattactt tcctttgcaa 180
acactagaac tcactgggct tgaaggtttt agtatagcat ttctttctcc aatattccta 240
acaattactc ctttctggaa attggttaac aagaagtgga tgctaaccct gctgaggata 300
atcactattg gcagcatagc ctccttccag gctccaaatg ccaaacttcg actgatggtt 360
cttgcgcttg gggtgtcttc ctcactgata gtgcaagctg tgacttggtg gtcaggaagt 420
catttgcaaa ggtacctcag aatttgggga ttcattttag gacagattgt tcttgttgtt 480
ctacgcatat ggtatacttc actaaaccca atctggagtt atcagatgtc caacaaagtg 540
atactgacat taagtgccat agccacactt gatcgtattg gcacagatgg tgactgcagt 600
aaacctgaag aaaagaagac tggtgaggta gccacgggga tggcctctag acccaactgg 660
ctgctggcag gggctgcttt tggtagcctt gtgttcctca cccactgggt tttcggagaa 720
gtctctctgt ttccagatgg gcagtgagtg ggcatccaca tccagggcca gatcctaacc 780
catttggagg tgcagtactg ctgtgcttgg caagtggatt gatgcttcca tcttgtttgt 840
ggtttcgtgg tactggtttg atctggtggg ttacaggaac agcttcagct gcggggctcc 900
tttacctgca cacatgggca gctgctgtgt ctggctgtgt cttcgctatc tttactgcat 960
ccatgtggcc ccaaacactt ggacacctta ttaactcagg gacaaaccct gggaaaacca 1020
tgaccattgc catgatattt tatcttctag aaatattttt ctgtgcctgg tgcacagctt 1080
ttaagtttgt cccaggaggt gtctacgcta gagaaagatc agatgtgctt ttggggacaa 1140
tgatgttaat tatcgggctg aatatgctat ttggtcctaa gaaaaacctt gacttgcttc 1200
ttcaaacaaa aaacagttct aaagtgcttt tcagaaagag tgaaaaatac atgaaacttt 1260
ttctgtggct gcttgttggt gtgggattgt tgggattagg actacggcat aaagcctatg 1320
agagaaaact gggcaaagtg gcaccaacca aagaggtctc tgctgccatc tggcctttca 1380
ggtttggata tgacaatgaa gggtggtcta gtctagaaag atcagctcac ctgctcaatg 1440
aaacaggtgc agatttcata acaattttgg agagtgatgc ttctaagccc tatatgggga 1500
acaatgactt aaccatgtgg ctaggggaaa agttgggttt ctatacagac tttggtccaa 1560
gcacaaggta tcacacttgg gggattatgg ctttgtcaag atacccaatt gtgaaatctg 1620
agcatcacct tcttccgtca ccagagggcg agatcgcacc agccatcaca ttgaccgtta 1680
acatttcggg caagctggtg gattttgtcg tgacacactt tgggaaccac gaagatgacc 1740
tcgacaggaa actgcaggct attgctgttt caaaactact gaaaagtagc tctaatcaag 1800
tgatatttct gggatatatc acttcagcac ctggctccag agattatcta cagctcactg 1860
aacatggcaa tgtgaaggat atcgacagca ctgatcatga cagatggtgt gaatacatta 1920
tgtatcgagg gctgatcagg ttgggttatg caagaatctc ccatgctgaa ctgagtgatt 1980
cagaaattca gatggcaaaa tttaggatcc ctgatgaccc cactaattat agagacaacc 2040

agaaagtggt catagaccac agagaagttt ctgagaaaat tcattttaat cccagatttg 2100
gatcctacaa agaaggacac aattatgaaa acaaccatca ttttcatatg aatactccca 2160
aatacttttt atgaaacatt taaaacaaga agttattggc tgggaaaatc taagaaaaaa 2220
agtatgtaag ataaaaagaa gagattaatg aaagtgggaa aatacacatg aagaacctca 2280
acttaaaaaa cacatggtat ctatgcagtg ggaaattacc tccatttgta aactatgttg 2340
cttaataaaa acatttctct aaaaaaaaaa aaaaaaaa 2378

<210> 96
<211> 1597
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3703614CB1

<400> 96
ttccagccct catgaacccc ttcttgggaa acttgccttc agcacccgcc atgggctgtt 60
cagatgcctc cactctgaac cctggaagtg cctcacatgt cagcacctat actgaggact 120
ctggatcagc acaccagagc agggatcaag tcttcctccc tgcattcccg gtacaagttc 180
gtaggtgtaa agctttgaag gaaaaagatt taattagaac gtctgagtca gactgttact 240
gctacaatca aaattcccaa gtggagtgga aatacatatg gtcaactatg caggtgaaaa 300
ttaccagtcc aggcctgttc agaattgtat atatcgcaga aagacataat tgccaatatc 360
cagaaaacat tctatctttt atcaaatgtg tgattcataa cttttggata ccaaaggaat 420
ctaacgaaat aaccataatc atcaatccat acagggagac tgtgtgcttc tctgtggagc 480
ctgtcaagaa gatatttaac tatatgatac atgtgaatcg aaacatcatg gatttcaaac 540
tcttccttgt gtttgtggca ggagtttttc ttttctttta tgcaaggacc ctgagtcaaa 600
gccctacttt ctattactcc tcgggaactg tgctaggtgt tctaatgaca ttagtctttg 660
tcttgctgtt ggtgaaaaga ttcattccga agtatagcac cttttgggct ctaatggttg 720
gttgttggtt tgcctcagtt tatattgtat gccagttgat ggaagatctg aagtggctgt 780
ggtatgaaaa caggatatat gtattaggct atgtcttgat agttggattt ttcagctttg 840
ttgtttgtta caagcatggg ccccttgcag acgacaggag cagaagtctt ctgatgtgga 900
tgctgcgact cctctccctg gttctggtct atgctggtgt ggctgtgcct cagtttgcct 960
atgcagccat aatcctcctc atgtcctcct ggagtctgca ctacccactg agagcatgca 1020
gttatatgag gtggaaaatg gagcagtggt ttacatcaaa agagctggtg gtgaaatatc 1080
ttacggaaga cgagtacagg gagcaagctg atgctgaaac gaacagtgct ctggaggagc 1140
tacgccgggc ctgccgaaaa cccgactttc cctcatggct ggtcgtctcc agactccaca 1200
ctcctagcaa ctgatgaaac caaggccctg aaaggttcac caagttgcca gggtcgtgca 1260
gatcaacaaa ctgcagaagc aagactcaac caggtcttct gctgtcaaat cctgtgccct 1320
tcccatggcc agcagagcct cccccagctg gatcagaaac ctttcaggga aatttagaag 1380
gtgctttctt tcatacacac aaaaagtagt gatcacttaa ccttaaaatc ttgtttcaca 1440
aagatgttgt aatattatac ctatacttaa aacgttagat ttctaataaa atttaattaa 1500
tagagctatg aataacttaa tagtaataat gaaagtcata gtaattaata gtaacttaat 1560
aaatagcaat aaaatagcaa tgaaaaaaaa aaaaaaa 1597

<210> 97
<211> 653
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4000975CB1

<400> 97
aaaacaaaaa aaaactgtgg gctctaccaa tagtcccaaa tgccaagaaa gcttccttca 60
tggagggaaa gccttttttct ctccgtggaa ctcagcccac tggccttggc aatgggcagt 120
gctcctggcc tgcaagtctt ctccaagaca aatcctctgt ttcttagccc acccttgaaa 180
agtagggctc tgggcccctc cccccaggag ggtttctggc ccaatctgca acgtcaagta 240
cgcgccgtgt ccttgggttg tgaggcagct ggggagggtg actttggaca gatgtccttg 300
ggctgtgagg cagctgggga gggtgacttt ggacagatgt ccttgggctg tgaggcagct 360
ggggagggtg actttggaca ggtgtccccg gcactgtgcc ccagtcaggt gcaactaagg 420
gatgggctgt gcctactttg aaggtaacag agtctgccca gggccacagt ccaagggcga 480

ggccctttg gggtgcctgt catggcatct ggtcatcagc caccttggcc cctgcaagct 540
ggggccaccc atgggcactc ctgagcaagg cctgggaggg acctgccagt gtcctgggag 600
ccactctgcc aaggctcagt ggcgatgtca tactttggaa aagaaaaaaa aaa 653

<210> 98
<211> 3090
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4598831CB1

<400> 98
ggctgggccg ggctgggcta cgcgcacggg ctcggccgcc gcccctgccg gttgcatacc 60
ctgtcctgag ggcgcggcac ggagtgcatg cgggccgctg ccatgacgac cgccatcttg 120
gagcgcctga gcaccctgtc ggtcagcggg cagcagctgc gccgcctgcc caagatcctg 180
gaggatgggc ttcccaagat gccttgcact gtcccagaaa cggatgtgcc ccagctcttc 240
cgggagcctt acatccgcac cggctaccgc cccacggggc acgagtggcg ctactacttc 300
ttcagcctct ttcagaaaca caacgaggtg gtcaacgtct ggacccattt actggcagcc 360
ctggccgtcc tcttgcgatt ctgggccttt gccgaggctg aggccttgcc atgggcgtct 420
acccactccc tgcctctgct cctcttcatc ctgtcgtcaa tcacttacct cacctgcagc 480
cttctggccc acctgctgca gtccaagtca gagctctccc actacacctt ctactttgtg 540
gactatgttg gcgtgagcgt ttaccaatat ggcagtgctt tggctcattt cttctacagc 600
tctgaccagg cctggtatga ccggttctgg cttttcttct tgccagcagc tgccttctgt 660
ggctggttat cttgtgctgg ctgttgctat gccaaatatc gttaccggag gccttatcca 720
gtcatgagga agatctgtca agtggtgcca gcaggtctgg cttttatcct agacatcagc 780
cctgtggcac accgtgtggc gctctgtcac ctggctggct gccaggagca agcagcctgg 840
taccacaccc tccagatcct cttcttcctg gttagcgctt atttcttctc ctgccccgtg 900
cctgagaagt acttcccggg ttcctgtgac atcgtgggcc atgggcatca gatcttccat 960
gcatttctgt ccatctgtac actctcccag ctggaggcca tcctcctgga ctaccagggg 1020
cggcaggaga tcttcctgca gcgccatgga cccctatctg tccacatggc ctgcctctcc 1080
ttcttcttcc tggctgcctg cagtgctgcc accgcagccc ttctgaggca caaagtcaag 1140
gccagactga ccaagaaaga ttcctgaggc tggcaagtgg ggcaacgtgt ggaggaagcc 1200
cctcataatt tggagaaaac ttgatacaat agaagctgac ttttaaggca ttggctttta 1260
aattaataca tatatccaag gatatgttat agctgcagtg tttgaaagcc aaaggattta 1320
agagttttgt tgttgttaat aaaaggaata ctccttttcc ttttggatca tagcttaaca 1380
aggcacagga agggaaggga tcttgactaa gattcatgag acattgaatt aaggagaatc 1440
atcttcatgc ctgaaaattt agcaaaattc cgactatggc ctccaggggc aattcctaaa 1500
agctgaatgg ataataaaat tggactggaa agtaagtagg tggctggtcc tcaccctgtt 1560
ggaatggcta tcctactatg ctgttctttg gtaatggaat aaattgaccc aaggaccgaa 1620
tttcatttgg atttcaaatt gtccagagtg gaaaagcctt caagatgaca tgatgaatta 1680
ctcagttcat ctgatttctg gtccctcctt tctcgacaac tataatacta acccttttct 1740
caggataact gtctacacct ggcagttttc tctgacgtgc tgttcactca catccctacc 1800
ttgcatggta atataaagga ctaggaagca gtcatacttc caggaaatgc ttggattcat 1860
gtggacattc aggaagctta ttctcatata atactaatct aaacagtact agaaattaca 1920
gtgccaagag ccaccaggag gcccagccaa taagcataga tactatatgg tatcatggga 1980
cccatctatt ttttaccagt ggactacagg attacttgag agttatcagg gctgcctaac 2040
agaccaggag atctgggggt tgcaccaggg aatcgccata tttgaccagc atgttttaaa 2100
agctcttggt aggattagtt ggttctaagg atccctctag ggacctcatt atttcaagag 2160
gaacccaaag tccagcctcc tacatagatg ctgccccacg aaggacccac aaaactaacc 2220
tagttcaggg ttctcaggca ggcagttctg cttcagctta gagcagaacc cataaaatac 2280
tcaagtactg ggataggcaa gcatgtgtgt ttactgtgga ttggtccctg aaggctcctt 2340
tgggtgagaa catgtgaacc aggcaccctg gtttgtttgg agcattgctg cccagaagct 2400
tctatgggat aggtggtgct tgggattgat gtgttgtggc catgcagccc tccctgagga 2460
ttgacttctg cactaatcca gtgaaggagg ctgtgtcaaa agaagggctc agaagccctc 2520
ttttcagagg caatgattcc tgtcagtatg aggtcccttа gttactaaaa agggacatga 2580
tttaactcca gtttgatgaa cctcctccga gtttacttta ttgtcttcaa atcttttgtt 2640
ttcttccttt ttgtgagatt tgtgggtttt gtgccttata aatggaaatg tatgaacaca 2700
atatatgctc atgtagaatt ttctgttctg ggttattggg ataagaaaaa atatatattg 2760
ctcttcaact agtgaatgaa agaaacttca gaaagctaga attgcttatc aatcaaaaga 2820
ctttctcaat ctattttggc cacaaacaaa catattcaac tgaagctttc caataatctt 2880
tatatcaaga aagcatgcgt cttgtcagct acattgtttt cttagatgga tttctcctgt 2940

taatcctcaa atatctgaac ttctgtgtta cccaagtgtc ttatacaagc ttctggtgtc 3000
taggacaaat ttatggcaaa taaaattagc aaaactgaaa aaaaaaaaaa aaaaaaaaaa 3060
aaaaattctc ggcgcaagaa ttagctggcc 3090

<210> 99
<211> 1274
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4992201CB1

<220>
<221> unsure
<222> 1191, 1250
<223> a, t, c, g, or other

<400> 99
cggcacgagc tggtcttgaa ctcccgacct caggtgatcc gcccgcttcc gccccccaaa 60
ggactgcgat tacatgcgtg agccaccgcg cccagcttcc aaaagtttta agcagagctc 120
agaggtctta agcacaggca catcagagga gcatttttga aatgctttcc agcttcctca 180
ataggaatgg aagccaaatt ccgaatttat gactcctttg aggaagtcga gagctgtaag 240
gaaagccagg aacaggggca agggagagat gcatcccgaa tgatcctgtg ccaattcttt 300
ctggaatctt tgatgtgatc tcagctgccc tttctatact tgacacagtg attgtggcac 360
ccactggtct agctgtggtc tacaaggaac ccccaaaagg aagggcacag tgagcagggg 420
catcggcctg agtgacaagg atttgagagg gcaggttgga tgcagggaga ggactggcca 480
aatgccatgt gtctggactt aggctgcctg gttcaaactg gacttcaccc tttttgactt 540
catgatctgg tacaagttat atgaaaatgc attgctcctt ttctagtctg taaaataata 600
atgaaatgtg cactaataac tgggagacta cgcagaggaa atgaaacaag ctgcatagac 660
cacagagctc agagcctggc ctttaggaag ccctcagtaa gggttcatga tgccatggtg 720
tctgtcatca tcctctttat cctcatcatc accttcataa tctttttatt gttcttagag 780
aatagtttag agggactcat tccctgctat catgggtgag atgtctatga aaaggacaac 840
cagtggggga ggaaagcaaa attttgaata agatttctga gaccgccccc ccccgaccat 900
aaccaagaac agaaactcca cagtctgctg agccgacagt ttgcacgttg gtctcctccc 960
atctgcccac cgcactctcc tgttttttc ctgaggatga ggatacaaaa caaggctccc 1020
aaccgtccct cagcactcac tgaactgccc ttcccctctg ctgggccatg accacggaga 1080
acaagtccac tgtcctccct gcgtggtgca cgatggaggc tcagactcca tcctcaaggc 1140
tggcaagaag acaggtgaga catgagcctc ctgatacagt gatggttctg ncacagatgg 1200
actcacatga ggtgagccaa tattacatca tgctggagtc agccaagctn tagcaatgca 1260
gtccacttag atca 1274

<210> 100
<211> 1514
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5441583CB1

<400> 100
aatctgcaaa gttcctgtga gcgctgtcat tttgtcactc tggtttttca gattcttccc 60
ctggaggctg gagtttccag gatgtcaaaa ttacctctgc ttgggtgagc tatttcaagc 120
agctgggata cctgtgtcac tcctgctgtc tgccagtgac tgcccaggtg tctgctggtt 180
cctccccagg agtagggagg aaccaggtgg gctggctggg atgggtggat atttaaagac 240
caggccttgg acgctgcagc acttctatct ctgcttgatg cctgctgcca cgtggctggt 300
cctcctcctc ctgctgtggc tgagccttgg ggtgaagaca ggcagctgct cccaacccca 360
gaacctttgc tgtcttggga cggatcacca ctgcaagagg ggaagttgct actgtgatga 420
attctgccat gtggcaccag actgccaccc agaccacagt gtcctctgca accctggtaa 480
ctcacataca ggcccgattc cacctacagc aaagctggat gcgatggctg gcagaggcaa 540
acccctttgcc tgcacttcag gccaaagccg ggatgtggcc tagatggttc ctaaggtccc 600
tgacaatcct gagatcttgc atcttgtcta tttcaggtca aaggtgccta catgctcctt 660

ctagctttgt ttccctgatg ttccttgcca cctgctactc ctctctgagc tacttttcca 720
ggttccacag ggagaggttc agctgtccgt ggtagacatg agggtagaga gtgaggtggt 780
tgggttccac ttacctttct atcatcttgc agtatggatg tctcttgact ggtaccgtgt 840
agacttttga gggcacacag gagtctgcag agagatactg gggacattga gcagcagcag 900
tggggttggg aggcagaaat gaggacagga acatttacct tgtgtttctc tcagcttctc 960
agatgaccaa gatggtgctg cagatggtgc tgaggatgga gaacccacca agccccgcta 1020
ggagccacct agactggatg cagagcatgg ttcagcgagt tctccaaacc agcctcccag 1080
gaatgccatt ctccatggct gtgaggagaa taaagaagag agcctgactc ctctcctgag 1140
gcccccttcc caccctgagc cagcaggatc cacggagcag aggtcatctg tccccagctt 1200
ggcccactga ggccagcatg gctgggccca ggatgcttgt ctctcagctc ccatcctgtg 1260
tacttccaca ttggtttaac cagaggaaaa ccgaaatcta caattgtcat aaacacattt 1320
aaatgtgcgt agaatcagcc atacaaattg tgaaacatac attgggctca tggtattatt 1380
cactgtttgg tggtggttag agtgactata gcaaacattt cttgaaagta gaaaataaga 1440
acccagcacc ccttgagcta agtaatgtgc cctctgtcct caatattcct tcctgggctc 1500
cacattactg ccct 1514

<210> 101
<211> 1380
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1639243CB1

<400> 101
tagtcatcga actgccactg ccctcctctg agccccatca tgctgggttg ttatggaatg 60
ggacagctct gtatctggga aagtcctcca gcctcacctt cttggctctt aagtgtaggc 120
tgttaccact tgcccagcct gggcctcctc tctccccacc ctttcacaag gcagctgccc 180
ttccgcactc actggcccat cccttctttc tcatcgtccc acccttccac ccccgtgcat 240
ggctgctgta ggtctggttt ttttgttttt gtttttttta agacggagtc tcactctgcc 300
gccaggctgg agtgcagtgg cagaatcttg gctcactgca acctctgcct cccaggttca 360
agcgattctc ctgcctcagc ctcccgagta gctgggacta caggcacgtg ccaccatatc 420
cagctaattt ttgtattttt agtagagatg gggtttcacc atgttggcca ggatctcttg 480
acctcgtgat ctgcccgcct cggcctccca aagtgctgtg attacaggcg tgagccaccg 540
tgcctagcca ggtcttgttt tgaaaacctc actgtgggag attcaggcat cctccctaag 600
ccagctggcc gctgtgctaa agcctgttca gagttaataa taatcattag ctgaatggtg 660
ctggggcctt tcagcttcag atctctaagc acttgcaggc tgagtcagtc agccctcacc 720
ttccccctcc ttcctgggct gcagagtgta acagaatggg aaggcactgt gggaaggaag 780
tcaggaatct tgctgctagc cacgccttgc agtgacttct cgtctgggag tgggcactga 840
gtcctctcag taaactaata agacttgcac ctgacaaagg tcaagatatg tagggaacac 900
agtgtatgct aggctgagac ctatggtggt ggcaggggtg gctgttgagc ctgaacttcc 960
agtactcctg cccttccttc tgtttacctg gcttggccta caggggcac ccctggtctt 1020
gatgcctcaa gcccagcatt tctgggtccc ctctgcaagc tcagagagca gggaggcttc 1080
tggtagtgct cttgatgctc ctgtgtctgg ttggcacaaa gatcctgtgt aacatgaaat 1140
gaaaggtgca tcagcttggg ggctgggaaa cctgcagtat gggtttactc cgtccctatc 1200
actggtgtgg ctgtgggcaa accacttatt gcctgaccta cctcacaggg atgttgtgag 1260
ggtttgatga gagaatgaat gttaatagga attggaaaat tcaaagcatt aaacacatgt 1320
aaacaggtgg taaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaagatcttt 1380

<210> 102
<211> 942
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1335166CB1

<400> 102
gccttggcct tccaaagtgc cgggattaca ggcttgagcc actgtgcctg gcctattatt 60
gttatttaag atggagtctt gctctgtcac cctggctgga gtgcagtggt gcaatcttgg 120
ctcactgcaa cctccacctc ctggattcaa gagattctcc tgcctcaacc tcctgagtag 180

ctgggattac aggcatgcac aaccacactg gctatttttt gtatctttaa cagagacagg 240
gtttcaccat gttggccagg ctggtctcga acttctgagc tcaagcgatc tgcctgcctt 300
ggcctcccaa agtgctggga ttacaggcgt gagtcactgt gcccgaccag ggagactatt 360
atgatggcta tgagtggaca atgccaagaa actgaagtac actccttatc aggagtttca 420
acaaacaact agtcaaaatc aaataactca aagtccatgc aataaagata gttcaatagc 480
atttgagtcc aattggtcag tctttgttca gggatgatgg tgattaaagg ccaaaccctc 540
cccattaaat gagagatcta cattagagag gtttcagaga cctccaccca gcaacagtgg 600
gcaagcagcg tttataaaca gaaaaaggaa gtggcattgg atataggctg attggttaca 660
gctcaacctt tgtctaattt ggacacatct ggtcagtctg cagtctgccc ctggctgaag 720
gctggctgct atgattggag aagactcctt tcttcttaca agaatatact gcctggtcaa 780
gttgtagttg gtttacataa taaataaggt tacatcccag tcaaatcctc cctacaaagt 840
atcccctatt ctgagctcta ttaccataga ctagtttttc tgctttaaat aaatgggatc 900
atgcaattaa cgaaaacaga gagaacaaat gcggcgccac tc 942

<210> 103
<211> 1815
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 166894CB1

<400> 103
gctcgaggtt gaggaggaaa gccattactt caccaatgtg aatcaaggtc attcgtcacc 60
tcgggaagca cacagagctg cgacgttcag gagggaggca cgctgctccc aagccaaacc 120
acgatgcagc ctggcaccag tcccatgcag cagggagaag gggagtttct ggaagcgagg 180
atgagctggg cagccaggcc cagggccctg gccgtgcatg catcagcccc tcctgagaaa 240
aagagctcaa ggactacatg gggtcttacc cacgagtggg gcaacggtgc agtttccagg 300
aggcgttggc aggagctgac cggccactga cgctggccgg ggtccctagg ggacagggac 360
tggagggtca gcggcctctg accgtccagc gtcatctgag gaaatcaggc caatcgtgag 420
ccaccagagt gtcctcaccg agacggcctc tctccctagt ataacaagcc acatttgctg 480
cctttctggg agagttagaa gcaagccacg tgggtgactt cttttcaaag ttctctgcct 540
cagaggagga agccccatgg gtggaggcca ggcccgaggc catcctgctc tgggccagag 600
cgtgcctgtg tctgatgagc gaggacatgc ctcgtctggc tgagccttcc atttcacacc 660
ggcagctcga acccccccatc acggcttcct tttgggctcc cctcctcagg ggaccgagct 720
tgaaatgtgc ttgcccttca tccctaaggc ctgcgtccct tcctcagaga ggactaagcc 780
cctctcccag tcttgatgtt ctccgaggcc cttcttattc acagaaccta tctggcttat 840
ttattcgctt gcttgcttct gatgtcttct cttactgaaa gcttgctgca aaggaccaca 900
cctgcttctc gacccaggaa cgtgggaaag ggaaaggctt ggcttgtttt ggtggagatg 960
gagatgctgg taacagtgga ggaatgccct ccttctgatt cacagtgggg aggtgctctg 1020
ggcccctgcc actgcccgag gacttcagct tttggttgtc ctgctgagag gatgcggcat 1080
ttgagcagta gcttctggag cccagagtga agccccaccc aggcaccctg acaggaccac 1140
aggaatgctc ctggcacagc aagcaggact cttgagaagc tcagcctcca cgctccttgt 1200
agatgttcag ttcaaactcc acagtttgtg tgactcactg aagggcttgg tttggctctc 1260
gctgacgagt ctgtcatcgg tgccaggagc tgagaacagc ccccttccta tttgaggctg 1320
gcctgtccat gcgcaccctt ggcctcacat caatggagga tcacccatcc cttcctcgtg 1380
ctagaaaccc aatggctgta ttccataagc ctgcaggact cttgctcttc tccttattta 1440
actacaccag cttaggagta gcctacatgc tccatcttca tttccttacc ccatctactc 1500
cacagtctac cattctcttg ctgaggttac ttacctggcc tctttcctct acactctttt 1560
ccacactcac ttgtccaggt gcacattaat acctacacct gactatccta cagataactc 1620
gaattcaaga tggtggggga ataaactcac catcttactc tatatgtctg cctgtccttc 1680
agtatttctc atgtgggata agggtgctat gattcaccca cttgtccaag acagaacacc 1740
tcgatgtcat tctttttta aatttttaat ttttttttga gacagagttt cgctgttatc 1800
acccaggctg gagtg 1815

<210> 104
<211> 1120
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 217969CB1

<400> 104
ctccgctgag ggaactggca ccttctccac acacctcctt cgccgagtca cagacaagct 60
atgcagagaa atctaccgga gccacggcaa ggctctggaa gaggacagag aagggagcag 120
cttccagagg ccacatcggt ttgtcgccat gacaacagag gggcggggcc gtgaccacct 180
gcccctggca caaggccccc aagtgccgaa gggtcactga gtcttaagca aagagcaaga 240
ggcatcagcc tcatgggaca atccacccag agtcgcctaa ccacagaccc caaaagtccc 300
ctagtggaca ccacccagtc caggaacaga tcaaattcaa gaatatgtaa acttgggtgg 360
gaaaaaaaat cttttcctcc actagccagg aactaatatt tagcattttc attatgaaga 420
taggcagcat accacatgag cagcctctgt gtgtcggtca ccagtaaaaa tcataatatg 480
ttcatggcac atgatggtta ttgcagtttt gttttttgtt tttttttga gacagagtct 540
gcctctgtca cccaacctgg agtgcagtgg taccatcaca gctcactgca acctcgacca 600
cccgggctcg agggatcctc ccacctcagc ctccaagttg ctcggactat aggcgtgtgc 660
caccacaccc agctaattct ttttagatgg ggtctcacta tgttgccctg gctagtctca 720
aacttccggg cttaagcaat tctcctacat cagcctccca aagtgctggg gttacaggca 780
tgagccacca cacccagcct gttgcagatg tcttgaaaga gtatttaccc ttcatctcta 840
ctgtgaaatt acggttgttt tcagacttgc cactattgaa tgcattatta aagtggtaca 900
tatatttcta catcgtaaat gtttatattt ggcatttga tatcagtatt tcaagagttc 960
tctccttcgt aatccaaagt gtgttatttt ttaaatgtaa acattattct aagaaggggc 1020
tcaggatcaa ccctggagga ccagacgggc tacagcacac acaaaaaagc ttcagtacac 1080
tgtattagtg tgctgaggat gtcgcaacaa agtacttcag 1120

<210> 105
<211> 535
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 335237CB1

<400> 105
aaatgaatag ggagcttctc tagactgggt agccccaaaa ggtctgccgg aagagatgct 60
atttagtctg aagcctggat cgtgatgaca gttccccctg tgaaggtgca ggtagagtgt 120
tttgtaaaca tgtgaccatc acctgtgatc atatctttaa ccctcacact ccccccatcct 180
aaccttcatg aaatgctgct catgctctga gctagtatct ctcatctaac ccctacccca 240
cctccaggag acctcgcctg ttacagtcta tgcctgattt agctgttgtc cttttctgca 300
gtagagtccc caggagttct tcagggacag gcagtcaagg acagctggtg cccagagcct 360
ccctggcttg cccactgggc agctccaggg acaacttgac ctgcccaatc aaagccaaag 420
gtcagaacag gaggcagaac ctcgcccggc cctcttcaaa ctccaaggga aagcctgttc 480
cttggatcct ttcagaaata aaaacaaaat gacatcaaaa aaaaaaaaaa aaaaa 535

<210> 106
<211> 1188
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 938306CB1

<220>
<221> unsure
<222> 1052
<223> a, t, c, g, or other

<400> 106
agtaactgtt ccttgcctcc ccggttcaaa tgataccctg cctcagcccc ccagatagct 60
gggattacgg gcacatgcca ccacacctgg ctaacttttg aatttttagt agagacagag 120
ttcgccatgt tagccaggct ggtctcgaac tcctgacctc agatgataca cccacctcag 180
cctcccaacg tgctgggatt acaggtgtga gccactgcgc ccagctggct tatgctcttt 240
ttctaaaaat ttgtccactt aggaaagtgt cttgatttca ataaaattag catggattaa 300

attacttttt ttctttttt ttttttttt ggtatttgat cttgctcgat gaatgtggac 360
atcagctatg ggttaggaat gctattcagg aacatctaat gtgtttgtga taatgtattt 420
gacttgtggc caaatggtac tcttgagaga tgtggttgtg actttgacag attaagccaa 480
agttaaattc tgatgaagtg caaaggaatt ttatcggtcc caggatggtt gccaactgtg 540
ttgggcaaaa gagtgatttt ccaaaaaggt ccagagcaat ccgcatgtat tctgtcacct 600
ctacttccag tttcttcaaa ggccagccaa aagcttcatt ttcctacttc ctgtcatttc 660
caaaaccaca gcttaaacct gaaaaacaaa tgggaagctg tgttcctgcc tcttatgatt 720
gctgccactt acaaaccagc aagaacagag cacagcaagc agaggagagt tcagtcttgc 780
tgatgtcaca tgaagcagaa tgaggactct gcttgtttgg ctcctgagtt taaaagataa 840
gtagacaatt tattgttggt aggttgagaa aattcaagaa cagagagaaa tctgacagtc 900
tggccataag aagaaatttt tgcatttatg ttcttatagc catacctgaa gtcaactaat 960
tacgtttgga gagaagtaga atgcttgctt taagaagaaa aaacaggcta gggcgcggtg 1020
gcacacacct gtaatcccag cactttggga gnccgaggcg ggcagatcat gaggtcagga 1080
gatcaagacc atcctggcta gcatggtgaa accctgtctc tactaaaaat acaaaaaatt 1140
agccgcgcgt ggtggtgggc acctgtggtc ccagctgctc gggtggct 1188

<210> 107
<211> 638
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1448129CB1

<400> 107
ctaatgtgga gaggaaggat tctgtggttc ccttaactgc aaaggattgt tcactatcgt 60
acatttagag ctgaagagaa ggcttgcgaa ggccatttga ctttttttc cttccttatc 120
gcagcaggtc catcagcaat gtcacgaacc tctaaccaag acccctcaag caatccaatt 180
tctggcagca cttaagtgct tttcatactc ctgatgagac atgatgctca tttctcaaca 240
aagatttatt gaatgtctgc tatgtttaat gcaccatggt ggagcctggg gaagatgcct 300
acgccatacc tgctgtcctt aatgaactca caggccagct tcgggcagac attccagcaa 360
gcattggagt ccaggctaat agtaaccaga gagaggtaca aactgggtga aagaaaggaa 420
cccttcctgg aggagtcagc atttgagcag ttcctgaagg tgctggtggg tagagggcat 480
tccaggcagg tgggtctctt cacagagtgg actgcagttt gggttgcctg acaaagatga 540
gaagggttag gtttttcctt cctccctcct gcatgaaata atgggaaaga agccgtgata 600
gaaagattta gtttggagta ggataaaggt gctaagaa 638

<210> 108
<211> 648
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1761049CB1

<400> 108
actgtaatgg gatactgata ctgcctcata gataatgaga ttaagtgtct cctaccttga 60
atgtctgctc cggtgttggc ctggagatat ctttcccaag tgtagaggaa tccatctctg 120
tgcccgtggt aatggctgga caccaggaga gtttgagatc agcctgggca acatggcaag 180
aacctgtccc aaatttaata taataaattt taaataaaat tttaaaaata aaattaaata 240
aaaaattata caggagggcc atgtcaggtg aattttaact gaagcaaacc ttatttattc 300
ataacaatga agtttatatt actaggaaaa aatttcattc catgataaat gtgtggtatc 360
atgtttttct tcaaaatatt gaatttaaag aatgttcatt gcaatactgg caactaagtc 420
cagatttgct ttttaaccat ggagtaattt ctgagaaata tttgttttat tttattttat 480
tttattttat tttatttatg ttatttatgt tatttatgtt atgttatgtt atgttatgtt 540
atgttatgtt atgttatgtt tttgaaacag agtcttgctc ttgttgccca 600
ggctggagtg caatggtgtg atctctgcct cctgggttca agcgattc 648

<210> 109
<211> 1181
<212> DNA

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1959587CB1

<400> 109
atttaaaatg tcttgttggc aaaatcatgt aacacctaaa gtcatctttt gtgcaaatct 60
catactttgg aaattactgt tccagggagt attcatgtga aaggaaagct gtttgaacta 120
tatatatacc caagcctgta agcagaaatg actggacaga aatacaagcc tgtgagacag 180
agagatctag gagatcccta gagaggtcca tctagacctc caggatgaaa ttacctcatt 240
tagctcagtt tctcacctct ccgttagtgc tttggtccac gggggtcagt ggttctgcag 300
gatttcacca actggttcca cagtgggaat gtgaagaggt acctggctgt gggaaaagct 360
gcttatcaaa gagaggcctg atagaaatgc tagggaaagt ggctgtgagc ctgcattatg 420
gacgggagca gagcggcagg gcatgttgct aggactaagc agtcgattca cgtgagtctg 480
gacaccatgt ataaatattt tgaaaacttc ctgttacaat ttaatttact ccattgtgtt 540
agagggagtg cttccctgca catggtgttc tgtatattta ttggttgatg tgctttaggg 600
tatagtaatt atttatagca atagattgat acccaggttg cagagttgct ttagttatat 660
ctttatttcc ttttcttaga gattatagac ataatttgct tctgtcatct tttgaccaaa 720
aatctaatgg gaatgaagaa ttacacttga atgtttattt ttaagtaata agttctgtga 780
tagtgattat aggtgtctgc atttaaaaaa atgttttaaa aagtacagag gtcaaacaag 840
catctgtatg attgaacaaa tgttatttcg tatatttctg tcaaatttta aaaaatggat 900
aaagaggccg ggcactgtgg ctcacgcctg taatcccagc actttgggag gccaaggtgg 960
gtggatcacc tgaggtcagg agttcaagac cagcctggcc aacatggtga aacctgtctc 1020
tactaaaaat ataaaaattg gccagatgtg gtggtgggtg cctgtaatcc cagccacccg 1080
ggaggctgag gttggagaat tgtttggacc caggagacgg aggttgctgt gagcccacat 1140
ggtgccaatg cactccagca atgggcgaca gagtgagaca c 1181

<210> 110
<211> 1291
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2303463CB1

<400> 110
ggcctctcac ggagactttc ccaccagtgt aataaagagg agaaaacgtc acagcggaag 60
ggcctgaccc tgctgcatcc actaaggaaa cagctacggg gatgggaccc tggaagctgc 120
tgtgggagcc tcatccaccg cttctctgac cccacccagg ctgcttccca ggcctcaggg 180
tcttagtgtg gacctccggg ccgtgattaa tgcagttttt caggttaagt actgcacgac 240
tactcgcttc tgaaactgat agacactgcc tcagctccgt gcagggcaga cgcacaagag 300
cagaatctcc gtgggacatc tctctggagc atcaatatta ctgcagtatt tggaagaaac 360
aaatttaaat aagttctaag gtaaagaatg gaacatttaa gacaagtctg gaaagtcatc 420
tgcctttaat aactgtcgtt tgtccttaac gtcagacttt ctccaagaca aaaactctaa 480
gaacttattt ccattcttac aaatagtaaa aatgataaat catatcaagt caattgaaag 540
tcctgcctgc tgctttccta aatcacaata tggccttggt atggttttat ttgtactttt 600
gtgggggatt cgtggatctt tagattaaaa aaaaaaatag gtttttgaac aatttttgca 660
agtttgcagc tgttagttat tgatttattt ttgcaaccca tctaattctt ctgtctctct 720
cctctcctca ggctcaataa ttccatgggt ctctaaggct gtgtttattt tctttaaatt 780
ttgttgattt acttttatac tccatttgat tttttctatc tctcaccttt tctcagactc 840
agtcatccca caggtctgta aggctctgtt cattttcttt aaacttttct ttttcctctc 900
ctcagattgg ataaattata ttgctatgtc tctgtttctg aactatagaa aagctcaaaa 960
ttactatttt tatttctcat tttttatatg gctatttttt ctctgctgat gtttcacatc 1020
tattcattta tgagaatatt ttcctttgcc ctcatgagcg tgtttataat agctgccttc 1080
aaattcttgt ctgccgttta catcttggac atcttggaga tggctactgc ctgcttttta 1140
tcttgtgtat ttattacatt ttcacgtgtc ttcacgcatc tcttgaattg gaaattgtgc 1200
cctggagact gtatacaaga ctggattaaa aagactggat tctgaaaaaa aaaaaaaaaa 1260
aaaaaaaaaa aaaaaaaaaa aaaaaaaatt g 1291

<210> 111
<211> 594

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2512281CB1

<400> 111
tgatctccca acttgcagtt tcaaggctgc ccttcccttc cctgtgctgt gggtttctcc 60
cccccgggct cccccgctcc tccctcagtg ctctgagagc acatggcagc tgccccagcc 120
ccgaagccaa gcctggcgcc tgtcctcggg cctctggaag tcctccccgc ccctctacaa 180
gcccctacca gacgctctcc agggacggag tgtgccctc cagccacagg aaaggggaga 240
ctgatccgcg tcaggtcccg ggatggaata gtaaccatga aatccagccg cagagcaatg 300
tgtctcaagc catcagtcac ccttcctaat tctcaggaag ccaggcatgc ccttcaccca 360
gctgaacctt gaactccgcc ttggcctttg caaggaaatc agccagtaaa taagacccca 420
aaatcggagc aacctgagaa gacatgacca gggtacaagg agggagaact gtggccaagg 480
atggccacag gccccagatg gctcacccag ctcaccaaag aactggcctc ttggaagagc 540
ttcaacattt tcagagatac cacgcgaggg gcggtcattg actgacttct ttcc 594

<210> 112
<211> 852
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2755924CB1

<400> 112
tgagtttatt tcaccacagt cgttaagcag gatgtgatgt ttattaaaat tcataattta 60
tttttctgta tctgtgtgct gcccaccttg gccattagcg gttggagctg cccctctttg 120
ttgtcactgt catttttaa acatagtata tgtattcttt ttctttttt agtgacaggg 180
tttcactatg ttgcccacac tggccatgaa ctcttgagct caggtgatct tcctacctca 240
gcctcccaag tagctgggac tacaggcaca tgccactgtg cccagcttgt cactgccaat 300
tttaatcttg gcatgtttgt cccattgctg tattgccatg tgaaaaattt tgccaacagc 360
caggagacct ctgtttcctc tgtaaagttg aatctttctt ccttgtaatg acttaatagt 420
gatttctatc aagtggagaa gaaatcttaa aaccgtacat actgttttgg gctctctggt 480
aggttttagg aggggagctc cccaggaggt aaatgggaat gccatccagc tttggcaaat 540
tcttattgaa actaatgagt ttaatagact gaccttatta ctcctattta caacctcatt 600
acaaaattat ctgctcaatg tttgctaaat tctttatcgg tcttgagaag ctaagccccc 660
tgtttaattc ctgctaaaac cgtgaaaatt ttcccccttt catatgaaaa gagtagcccc 720
aaaggcccct gagtaaaaat aaatagtggc tttggcaatt ttgacatcag gaataaccta 780
gggaaatttg gcttaaagaa tttaaagcta aaaagtggtc tcctatacca cgtgaataaa 840
aaaaaaaaaa aa 852

<210> 113
<211> 1361
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2796369CB1

<400> 113
aactgtgtgt acacagtgag cgtaagtgac tcagaagact caaaagatat ctaacccatg 60
ctaagtggaa atttgggccc gcgtaaaaat ttcaccacgg gcaacagcta tgacccagga 120
ttccccccaag gtcataatac ggccccctat tgggcaaact ggtaccccaa aaagtcccgc 180
tacgggaata attcccgggg agacgagtcc actagtcagg ccgctgccct ttttttttt 240
ttttcaatgt ttaagaaaat ttatagattt gtgttgggct gcattcaaag ctgttccatg 300
ggccaacagt tggacaagtt tgccctagat catggggtct ttgtgctctg agaatgggcg 360
agtttgggat ggactcagct ttctgttggt tggccctgga tctggatctg gtgctgcccc 420
ctttctctgg agcacacaaa gggaacaaga aggccttgac ttgggcaagg aagccatcca 480

cagggctccc cagaagccag ggccaccagg agctcactgc tgtgcagagg caactcggct 540
tggttatttc ttgcctgagg cagggaacag ggaatgcagg gaagccaggc agcagcagga 600
ggctccaaac gctggtgtca gcaagccaga acccccacct gacttcactc ctgtctgccc 660
agctcacagt cgcctgtcac tgggaggtcc ttggggcctg gacctgcctg acctgtggcc 720
ccagaagggg ctaagcccag agtcacatgg gatggagcct ggaatgcata ggccaagtgg 780
gttgtgcctg ggcagcaggc ctgggatatg aggggcctac tcctccatcc acagccccga 840
cacacattct ggagggatgg tcgtacctct gtgaccagcc ctccgcaccg tcacataggc 900
cagcaggggg caatgccagg tgccatccag agcccacccc catgtcctcc cggaggtcat 960
gagctctgcc ctacgtaccc cagccccaca atgggaagga ggcatggagc tggcagagca 1020
tgcgccacgc cgtccctgtg tgtgtttctg aggaactcct gttcactctg aggccactgc 1080
ctaggtgccc gtccctggtc tgctgctgct ggctctgatg aagctaacca gccttcttcc 1140
ccaccaagag ggtgtcccag cagcagctat ggctatgttc cggcgggggg aaggcgggga 1200
gccttcctcc tcactggctg tttcattttt ttatttttt gaggcagagt ctctcgtcac 1260
ccaggctgga gtgcagcggc gtaatctcgt ctcgctgcaa cctccacctc ccaggttcaa 1320
gcgattctct tgcctcagcc ttccaagtag ctgggatcac a 1361

<210> 114
<211> 1650
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3010920CB1

<400> 114
ctttttttt ttggtgaaac agccaagtga ataaaactat ctagcctagg tgaggtggcc 60
tgtaatccca gcactttggg aggcccaggc gggggatca cctgaggtca ggagttcgag 120
accagccagg caaacatggt gaaacctcat ctctactaaa aatacaaaaa ttagccaggc 180
atggtaatcc caggtacgtg ggaggctgag gcaggagaat tgcttgaacc tgggaggcag 240
aggttgcagt gagccaaaat tatgccattg cactccagcc tgggcaacaa agcatgactc 300
tgtctccaga gggggaaaaa aaaacctgtc tacaaacttc tagtaagtta tttatttcat 360
catatggcac taagaaaatc atcatgtttg ccattaaaac taggtacact aataacttat 420
agcttaatat ttttggcatg gttcttactg aaatctgcta ctttcaacca agtcattatg 480
cctcgagagc tctgtcaaga cttaatatat gttcattcct atgataaata tttactgata 540
tttcagataa atagttgtgg ttgctgtaac acatatattc actacagaaa actgtagatt 600
tctatataag caagaaccaa ttattgtatg atcataagca ataccataag ggcaatatga 660
atatttacta gtattataca aaacagctag caccaaatat ttggatttca gtacagttct 720
tcctccagta cccacatcct tgagacgtat taggaggtaa aattcaccta agttttagaa 780
ctcaacaaac catatctcaa gtatttactt ctgcctttac catacgtgtg atcctgggca 840
agttagttaa cctccctcgg ccttagtttc ttcatttgta aaataataat acaaaccttg 900
ttaagttttt aacagaacaa tatattgtat gtaaagtacc taggaaaggg tctggcccat 960
agtaaatgct tgataatgac agtgattact gtaatcattg ttgaaattgc agagatcaaa 1020
caattcctaa taatatagcc aaaaaagtga agtgctcagg accttcaaaa atatttcaaa 1080
gaatttgtta tcaagatatt tttagacttg tatttctggg ttctacaaag tccagatggt 1140
atcttctgtc ctacacaagt taggctgtgt cgtaagactc ttctaccagg tactttcttg 1200
acatctgtga gaggcaagca ggtatttccc cagtacttaa acactaggtc tgctaattac 1260
tgagtcagat ttaacgaata cgtaaactac tctggctatg ctaaacactt agggaggctg 1320
catcagattt ttcaacaact ggtcagtatg tactgagtcc actgagtacg cagctatgtt 1380
ctaaaaagtt gcattcacat caataatcaa acataatccc tgcctgagta gagtttacca 1440
tctagaggaa acaaactaaa agaagtttaa aagaaaatgt agtatgtttt agtataatgg 1500
ctagctaaag ttgattttgc tgtgttttaa aaacgagaaa gaatataatt agatgagagg 1560
cagttataaa tataaattac aatttggatg ttgaatgtag caacagcgag aaagcacacg 1620
aacacaggat gaaacataag taatgctgac 1650

<210> 115
<211> 1845
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3360955CB1

<400> 115
gaccgcagga attttttttt ttttttttg cattgtattt tcaactttat taaataagca 60
aaaaaaaaca actggataaa aagtaatttt gcaaagtcta ttaagctccc atttgtttgt 120
ggaggtgttg gctagttgtc tagatccttg agtaactgac ctttttagtt tgtttttatg 180
ttgtttgttt tcagcttttg tccacaacag gcagtcactt cagaccagga agtctctaaa 240
agcaccgaga ctcttcggag gttaatgctt tcggccaaga tcatggatgg ggaggacaca 300
gggctttacc atcagcactt ctcttggtac ctcactatca acaggatgat ggcccatagg 360
agcaaaggca ccagttttca tgcactgccc agtttaccca ttttggcaaa tccttcctcc 420
tggcctcctg actatgacac cacccagatg tccatatttt ctgccaggaa aagcctgctg 480
gggacaaagc ttcttacttc atgtttatct tctctccatt tcagaaaatg tcccgtctta 540
cattgcaact tgttgaaagc aggaaagtga ttttatagta agactaattt acattaaagt 600
gaaaatgaca ttattttgag gggtaccact tcaagtaatt tgtttatatt ttaataatat 660
attgagagtc ttggtggctc aggcctgttt tcctggcaca taaggaggca aagctgggtg 720
tttgagacca gcctggacaa tataggaagg cctcatctat ataacaacaa caacaacaaa 780
aaagaatatg atagagacat tcattgagca ttacattttt tgagcattat atattctgtg 840
tcctgtgtct ccatccaaat tttagattca aaatatatta ttactgtaat gttaacccca 900
aaagaaaaat atttattggt aataatctaa gtactatgat atgtaaagaa gacctctcca 960
aatgctcaaa tcataccata ttgggaattg attactcaaa tagctctctc tatataccta 1020
tatctaaacc aggaccaaag actgaacagg ctcggaagat atctttgttt ttatttttgt 1080
tttagaaaca gggtctctct ttgttgccca gactagtctc gaactcctgg gctcaaaaga 1140
tcctccccac tcaggctccc aaagttctgg gattacaggt gtgagccact gtgcccagac 1200
tgaagttatc tttgtagggt gcttctttca catgggagat ggaacataca gatccaatac 1260
agatgggatt gttctataga gtaaaacgaa caagtgcata attaattccc atgtaagcca 1320
ttctctggtc aaaagtttgc caaaagtttc taccatttga tcctctatct ctgaagtgaa 1380
tacagaagca gtacaatgag gcagctatta gcagaggaag ggacctacat gggactcttg 1440
tagctcatgt ggatatggaa ctgcagaaga gagaggctga ggttgacgta aaataagacc 1500
ctgaaatcca aaggcaatca acgagatagc attggttaga gttcaaaacc tcagagatcc 1560
atatggcttt cttggggaaa aactcattta tcttgaccta aggcagaaag tgagaaggag 1620
gcagtagcca ccaagacaga agaaaactta agcatttgaa gagggaaatt caatgggctc 1680
caggtgcctt ccatgtaaac atttcagaaa aactcccaga agtttgatag agctcatgcg 1740
tgtcaatgaa actgtaaaat ccagtcaagt cgccatcctg gttggactca tagagggaac 1800
ggaggagcca agatggaaat gcagaaatca cccgtcttct gcgtc 1845

<210> 116
<211> 1061
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3409459CB1

<400> 116
agctgacaga caagctcgag cacttagagc aagagaaaca tgaattgaga agacgatttg 60
agaaccgaga aggggagtgg gaaggccgag tgtcagagct ggagagtgat gtgaagcagc 120
cgcccctcct cctccaccca ggccccaaat caaaggctcc gcggcccggc cagccctcag 180
ctgctcacaa ccgattcagt ctccctccct ccctcacgtg gggaaagcac agcagggatg 240
cgcggcaaga atgtacctgt agatgtgtac ataccacagt gctgtaattt tgtatgtagc 300
aatcatgtaa atacatgtat ggattttata atatacatat ataaaaatct ataaaggcat 360
atttttagaa aaacagcaca ccactgcttc ttttgaaaat agtctgaata agaataaaat 420
gaatttctac agagcttcct gcctcagtct ctgggtattt gcgggggcg ggtttggtct 480
gaatgcagct gacatgtcag actcaccact ggctgctgct ggagaggtag caatagtggt 540
tcccctgcac ccgggccacc tcagatgctg gtacttactt aaccaaggga tttggccagg 600
aagggcatcc tcccctgccc ctcccgcctg gcactgcccc ctgccagtgc tgcagcgtgc 660
catccgcaag gcaggcctgc caacgctgct tcccaggcca gcagggccct aaagcttggc 720
agaaacttca agccagaaag aaaccacgcc aggcaaaggc ttccactgca tccttccagg 780
catgtggagg ataaacacgg agcttccccc ggcatcttac cacatgtgct cccagtccaa 840
ctaggggggag actcaaccct ggcccacctg ttcatggaga aaatggacct ggcttctgga 900
agtgtcccaa gggcctcgac aggctcctgt gcaagactta ggccaagaac agggaagagc 960
aggacacggt gggtgggaag aagccacggc cactcagcac atctctggag ggagagcaaa 1020
cagaccctgg aagaaaattt ggcagcttga acctttattt t 1061

<210> 117

<211> 1085
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4102938CB1

<400> 117
cctttcttga tctccgtaca cacgcatacc acccaccctc cctccctctc ttctcccact 60
tcccctctct cctttctcac catctctcct cattatgcct tctctactcg atcacccctt 120
tgcagagaaa cctttcctcc ttcttgctct tttccagttg aatttcctcg ccccgttgtc 180
ccaggtggct ggacatgcag cagaagggaa ctggggagac tctcgcactg ccaatcattt 240
ctcaaaactc aggttccagt ttgagacccg cctggccaac atggtgaaac ctcgtctcta 300
caaaaaatac aaaaattagc cgggcaaggt agcgcacctg tggtcccagg tactcaggag 360
actgaggcat gagaatccct tgaaactggg aggcggaagt tgcagtgagc tgtgatcgtg 420
ccactgcact ccagcctggt tgacagagca acatgctatc tctaaaacaa acaaacaaac 480
aaaaactcag gttcccacac cctctaaacc ctgcctcctc tcaggctaca gagacctctc 540
caggaggctg aagtgccctt accccgacca tctgaccagc caccgcccca tgcccgtgcc 600
ccaccgaggg cggaggctgc tcactgctct gttttatctc tggcctctga tcctgcattc 660
ttgtgccagg gcttagaccc aggcaaactg cttccggtag tcaataaacc ggtaaaccag 720
caatagacat aagcggctat ttaacgaccc tgccctgaac cgacgaccgg gtcgaatttg 780
ctttcgaatt tctgccattc atccgcttat tatcacttat tcaggcgtag caaccaggcg 840
tttaagggca ccaataactg ccttaaaaaa attacgcccc gcctgcactc atcgcagtac 900
tgttgtattc attagcattc tggcgacatg gaagccatca caaacggcat gtgaacctga 960
tcgccagcgg atcgactgcg gtgtgggtta atttttggcc tggggaacgg gggcaaaatg 1020
tccctttggc cgcttatcca ctggactccc ccggtggtgg ggaattttct acttggaaag 1080
cggta 1085

<210> 118
<211> 870
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4124601CB1

<400> 118
ccccttttt ttttttttga gacacagtat cgctctgttg cccagactgg agtgcggtgg 60
cgtgatctat gctcactgca acctctacct cccgggttca taccattctc ctgcctcagc 120
ctcccgagta gctgggacta caggcacccg ccaccacgcc cggctaatct tgtttttgta 180
tttttagtag agacggggtt tcaccgtgtt agccaggatg gtctcaattt cctgacctcg 240
tgatctgccc gcctcggtct cccaaagtgc tgggattaca ggctattaca cgtcctttta 300
tactgactgg acttttgatc taaaaaacag tttggctgac ttggagtgat ggagttacta 360
gtctctgtgt cagggactgc aaatgctggg cctccagcag gggcagggcg gatccagcga 420
aagacagaag tgggtgggcc caaggggatg gagagctgca gaacacaagt cccgcctcaa 480
gggggcagca actgcccaga gtccgctgac ggccgctggg tgggactgca agcccagggt 540
tgccagatct gtcagctttt ttcaagataa gcttgaaatc agattttcac atggaattgt 600
cagctaattc aacaatttaa aacacagcat gggctgaatg caacatgcct atggttcaac 660
attggctgca ggccaccagt ttgtgcgttc ggatccagga ggattccatt aaggattgtg 720
cacctcagag ccatctgcaa agctggtgca gacccttcat ttgcttctag aaaatcgggt 780
ttactccaca tggggtccgc aaaggatgag gacaaagcca gtgccctgag gtagaccgtg 840
ggagtcaccc acccactcat cattatttaa 870

<210> 119
<211> 3394
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4180577CB1

<400> 119
acgagccgga ccactagtta cggcgcagtg tgctggaaag attaaagagg tcgctgttaa 60
atgtaaagaa aaaatgctcg tagctatttg cttcctggta ttggagcagt tcagttgttt 120
agtttatacc attggattca attcattgca ccatggttgc caaaagtgcc tgaggtcata 180
atggattgtt aaaataacta aattccagtg gttggaaact ctaggtttgt accatttttt 240
ctgctgtggg aaaaaacaac aacaacaagc atgatcaagg taacatcaca tttgatgtat 300
aatattatac tattaatgga atatcagtag acaactgtta acccattagt agcatgagta 360
taaacagtac acctgaataa attggagaca ttagccacta ggtttaacag tggaatcttg 420
atttgcctag gtgacttctg ggattactgt ttgacaaata agagtaacat tttatttcat 480
ttcagaattt acgtcacttt tagctacaag agtaggaaga aggtaatcgg caaggcagaa 540
gagtatactc tttgccttag gatagcgtaa actcaggctg agacataccc ggcttataga 600
gttcttctag atgtgtagac tgtaaatgcc caaatcctct caactaaagt tttagtgatt 660
ccacaaagcc tctcatgtaa atttccagtg attccaccat tgcacttgtg aatatgtatc 720
cttgttagta cccagggatg tcctcgagca ccagttttat tttatctgcc attgcatctg 780
gattccatta cagcctctca gctgttactg cctgtggaca gttacttctg cttactgcct 840
gtagagagtt acctaacttc tcttctcagt tcttcctcag gtcctggcta ttttggcctc 900
agttgaaggg agtcttgctc tcatctctga gggttttaag tttgtttgat cccattgttg 960
tcttttctag ctttgagcat gtttttcagt attcatattt taacttactg agaacattaa 1020
agggaaatga taaactcgtg gtggggatat ggcagacagg tgcttgtttg tttgagagaa 1080
gtagcagaag agataaaata caaagtgcta tatgtttcag ctggagagga aagagagaga 1140
atttattaga ttatatactt gtcccatggc ataccacgta tatgtttaaa tagggacaca 1200
tctccctatg tttaactata cttataaaca actttgatac acattgcgtc ttttattctg 1260
tcacctgata ttttagtgta tctcaagtta cagattacat gtgtccttaa actatttctg 1320
aatttggact tagttccata tacagaaaga actttagaaa attcattaat ttggatcttc 1380
tattgatagc cataaatatt atgtttatgt attctaaaac ctctttgttt agttagtact 1440
gttcatgaat gtaacaagct tcaatttctc atttgtgagt agtacatttg ctttttgttt 1500
gtttgtttgt ttgtttttga gatggagtct cacgctgtca ccaggctgga gtgcagtggc 1560
gcgatttcag ctcactgcaa cctccacctc ccaggtgcaa gtgatgcccc tgcctcagcc 1620
tcccgagtag ctgggactac agacacccgc caccacacct ggctaatttt tgtatttta 1680
gtagagacgg ggtttcacca tgttggctag gctggtctca aactcttgac ctcgtgattt 1740
gcccgcctct gcctccccaa agtgctggga ttacaggcgt gagccaccac gcccagccgt 1800
acatttactt tttaaagcag cagactaggt acactaattc tcactcaaat attttcatgg 1860
gaatgtagtt atcaccaagt cctaaagtat tatttatgcc aaaaaaaatt tcattttaag 1920
gactacaaaa atgattctaa ttaaacattt tataatcaat agtaggttgg gtctttagcc 1980
attatatgtg tatatataca gacacatatg tatacactta cattttgaca gggtcttcat 2040
tgagtcttga tgcactttaa acccagctgg ctaccagaga tgcgaaggtg ggctctttga 2100
agattagcaa aatggacgtt tctgtcactt gagaaaagga aagttctttg cctttaaatt 2160
acacagtttt catcatgccc acaatctata ttattggctg gttaaacagc actgccctat 2220
tagcaatgtt aacaaaaatg aaattattta ttggcggtta tagattatct aattcaggaa 2280
atttctgagc tcaactttta cagcaactgt tatgccttct aatttagcaa ttgagttatg 2340
agtaagtttt gtgcttaact cctagaccct attgttgata accagatcaa atatagtctg 2400
tacagaggaa aacactggga acatttagta tttctaaagc ctcctttgga gttactactg 2460
attgtaattt ggaactgata ataggtagag attgctaaca ctgtttttt tcctggatct 2520
tttttatgcc agaaattaaa caggttctgc taactctttt ttttctcttg gttatcacca 2580
gaatgaaaat atttaaagtg atgactctag aaaagccatc tgtgcctggt taacattgag 2640
tttgagtctc ttcaatatat attgatcatg tattgattaa tctttatttt ttcatatttt 2700
ggctagacaa attcagatct atataatgga atacccctte ttgagtgaac tatactacta 2760
atctacatga ttatatagta aggaaaaaag aagaaataac tgtaataggc atagtgtttg 2820
ttgttggttg tcttgtcatt catgtgatac tactcatttc caaaattcac acaaacttac 2880
atgaggtgga ttatttgttt tgttcattat ttagttccta tatgtttttt ctttagaaac 2940
agagtctcat tctgtcaccc aggctggagt ccaatggggc ggtcatagtt cactgcagcc 3000
ttgaactctt tggctcatgt gatcctccca tctcagcctc ccacagcagg tgagactaca 3060
ggtacatgcc actgtgcctg acttttaat ttttttgtag agacgaggtt tcagtttgtt 3120
gcccaagctg atcttgaact cctggtttca agcgatcctc ccacctcggc ctcccaaagt 3180
agtgggattt caggcatgac cacctggcct agttcctata cttttcttaa ttcttcagac 3240
ttctcacatt tagtatagtg cattcatttc atcttgctgt ttattagcac cctttgtggc 3300
cagggaaata aaaggtggta aaattcagtt ttcagtttag ttcttgaaag ctctgggaaa 3360
tggagaactt tccaagcaca gtgtcgataa tcaa 3394

<210> 120
<211> 2343
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5265807CB1

<400> 120
gatgtggagg aagtgcgttc cctactctcg cagaaggaga acatcaatgt gctggtgagt 60
tgggagcagg gaggatggcg agaccggaat cctggctccc gtgttctgga cccctcccca 120
ccggtggggc ggctccgtgt ctcctctgct cagatgaggc cgcagttctt tccatcctgg 180
agtgaaagcc tactctgcat ttggggatgg agagcaggtc aggcttgaag cccagatcac 240
ttgggctggg aatgcccagg aacgttgttc tcttctcaac ctccggagtc tgcagcgccc 300
ccacccctcc tggaaaaagg agccggcttc tcagcccatt ggccctgact ccacccctcc 360
taggtgcttg cctttgactc ccaataccgg gacttaggac gctgcagaaa ctatgggaag 420
aggttgtggg gaagggcctg gagctctgag acttgggctg gggtccttga tagccccctt 480
accccgcgcc catactccct gctctttttt gccttcccag gaccaagaga ggcgaactcc 540
attgcatgct gctgcctacg taggcgatgt ccccatcctc cagttgctac tgatgtcagg 600
tgctaatgtc aatgctaagg acacactgtg gctgacccct cttcatcgtg ctgctgcctc 660
ccgaaacgag gtatgcaccc ccaccttcta ttacccttcc aaaaaggata agttgcttaa 720
aggtggaggt taagggttaa ggagcctcag tcctcagccc tacaagctgt tctaccccag 780
gtagtgagga gatcaaatct tggccccttc ctcctacccc tagaactggc cagtaagggg 840
gttgagtgct gccagctgat gtgactcagg aagagcttaa cactgggctg gcagactgcg 900
ggggccctgc tgccaggcct gagtgactca tccctccctg gggggccacc ttgagtctga 960
tcccccctctg ggcactgaga tttcttttcc ttgtgatgga ggatgggtat agtaatctgt 1020
tgcatccttg gccctcaagc agcttgaaag ttgacagcac aggggccaaa catatttcat 1080
cagttgggtc ctttatccta ttcccatcag ccctgaagac agggaaaggg ggacctcctc 1140
ctcagttcat cacttccctg gctccagcca ccaggggggag gggccgaggc cagtgggaat 1200
gaggtgggtg gtggttagga ggggcatttt agctcgggtt ttgcaaacga accagtatct 1260
gtttgcttac tgagatactg aaaataattt tggtgcctat gagaaggcta aagagggtgt 1320
tcagggcaag atgtagcaag agcatttagt tttcctagaa gtagtagggg gttgggatgt 1380
gtggatctcc acatgtggat ttgtttggga atggatgtat gtgtgagtga agattttcag 1440
ggtctaagtt ggcgagtcag gtggcttgaa gaagagcaaa gggtccttgt tttaatcagc 1500
aaggtaagga agagcttaag ttctaagaag atccagaaga tggaaagtag gtattcagag 1560
gctaggtatg ttggaaaaat gagaagagat tctgtgctgg ggaaggggct gtagatgcta 1620
tgctgaccta ttcctccttc catttcctgt tgttttacct cctgctccca ctctctcttt 1680
tgtcccctgc ccctcaacag aaggtgctgg ggctgctgct ggcacattca gcagatgtga 1740
atgcccggga caagctgtgg cagacaccac tgcatgtggc tgctgccaac cgggccacca 1800
agtgtgctga ggctctggca cccctgttga gcagcctcaa cgtggctgac aggagcgggc 1860
gcagtgccct gcaccatgca gtgcatagtg ggcatcttga ggtgaggact gttcccatcc 1920
aggcccagct ggggctttcc cttttcttac cttcctactc acgtttccct gcttctgggc 1980
cctcaagcct gaaggagaag cagcccggat ggttatataa acacttaagc tgaagcacta 2040
aatgaatggt cagactagat taaaattctt aacttggact tcaggaggtt gataaaacct 2100
tggattttgc gtggaaaata tttgcataca cgcctttctg aagagagtcc atgtctttca 2160
gcacatttgc aaagacttac aagagtttaa gaaccggccg ggcacggcgg ctcacgcctg 2220
taatcccagc actttgggag gcgaaggcgg gtggatcatg agatcaggag atcgagatca 2280
tcctggctaa catggtgaaa ccctgtctct actaaaaata caaaaaatta gccgggtgtg 2340
gtg 2343

<210> 121
<211> 751
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5405979CB1

<400> 121
cacggtttga ctcaagtaac tttatttttc tctaaattat tttatctggc aacatttctg 60
tggctttaac tattaactgt aaacataaaa cagatgtgag gatgatctta ttccaagagg 120
tggccgctgt gttccagccc caggttgtct ggtactttcc ttcctcactc ctcacctccc 180
acgtggcact atgcaaagct tcacattcta tcttgttctt ccttctccag tcgttcttgc 240
tcctcctgtt ccttctgcag ccgggcctgt cttctctttt cagccaagat cttcgcagcc 300
tctcctgcat cagtggtgcc tgctgtgggc ttccccaagg ctgcgctgtt ttcggctttc 360
cctcctgcgg cagcagcatg cttctcgctg gcatgcttgt ccactacatg cttctctagg 420

gcttcctctc cttgcgggcc agcagcctgc tgagccaagg taccttccct ttccttgttg 480
ctcttctctt tctccgcttt cctcttgggc atgtcttggc cagaaagggc aggtaagcgg 540
tacttcacag gagaccctgg gtagggaggt ttcgttgtct ttggagactg tggataagct 600
ttagtagctg tcttggatat cacaggagaa gatggtctct taggaatgcc tccagaaaac 660
tcacatcttc tcagagggga cccgaagttc acaacaggaa gagaagttgc tgttcgttgc 720
ccccgcttca ccttctccac cagagaggcc t 751

<210> 122
<211> 618
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7481109CB1

<400> 122
atgatgagaa ctctcatcac cacacaccca ctgcccctgc ttctattgcc gcagcagctg 60
ctgcagctgg tgcagtttca agaggtggat acagattttg atttcccaga agaagataaa 120
aaagaagaat ttgaagagtg tttggaaaaa ttttttagta cagggcccgc cagaccacct 180
accaaagaaa aagtcaaaag acgtgtcctt attgaacctg gaatgccact aaatcatata 240
gagtactgta accatgaaat catgggaaaa aatgtttact acaaacaccg ttgggtggca 300
gaacattact tccttcttat gcaatatgac gagctccaaa aaatctgtta caacagattt 360
gtgccatgta agaatggaat taggaaatgt aacaggagca aaggtcttgt agaaggagtg 420
tattgtaatt taacagaagc atttgaaata ccagcgtgta aatacgaatc actttatagg 480
aagggctacg tccttatcac ttgttcatgg caaaatgaaa tgcaaaaacg tattcctcat 540
actataaatg atctcgtgga gccacctgaa cacagaagtt tcctcagtga ggatggtgtc 600
tttgtcatat cgccctag 618

<210> 123
<211> 979
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6247114CB1

<400> 123
gacctagagc aggcatgggt gggtcacagg ctttggagag cactctctgt cctgatcttt 60
tcagttgaga gacttcagct gttcattgct catttggact tagttcaaga attttgggtg 120
tcaccaggta aacagagccc tcagcatctg aatagaaact gaacaggaac agaagagatt 180
acactacatc tgagatggag acctttcctc tgctgctgct cagcctgggc ctggttcttg 240
cagaagcatc agaaagcaca atgaagataa ttaaagaaga atttacagac gaagagatgc 300
aatatgacat ggcaaaaagt ggccaagaaa aacagaccat tgagatatta atgaacccga 360
tcctgttagt taaaaatacc agcctcagca tgtccaagga tgatatgtct tccacattac 420
tgacattcag aagtttacat tataatgacc ccaagggaaa cagttcgggt aatgacaaag 480
agtgttgcaa tgacatgaca gtctggagaa aagtttcaga agcaaacgga tcgtgcaagt 540
ggagcaataa cttcatccgc agctccacag aagtgatgcg cagggtccac agggccccca 600
gctgcaagtt tgtacagaat cctggcataa gctgctgtga gagcctagaa ctggaaaata 660
cagtgtgcca gttcactaca ggcaaacaat tccccaggtg ccaataccat agtgttacct 720
cattagagaa gatattgaca gtgctgacag gtcattctct gatgagctgg ttagtttgtg 780
gctctaagtt gtaaatccca cagagcttta ggactagggt cttactaagg aaggacctct 840
tcttgttcat tcttgtttaa acctttcctt aatatctact ctttagcact atagtgaact 900
cctgattatt tattctaact ggaggagtga aaaatccaaa attgtggata attcaattaa 960
aagttatgac tgataaaaa 979

<210> 124
<211> 3012
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 3243866CB1

<400> 124
gccctgcctc tgcccgctcc ccgggcgcgg agccgcgggt ttcatggggc gattgcagcg 60
attcccccac ccagagcgac ctgcgggcag cggcggcagt ggcaggagcc gcctttccga 120
ttccctacga tgcgggtgct gagctatggc aaagggcagc gaagtgacga gcgagacccg 180
cgtacgactg tgaaagccac ctggagccac cttgccggga ttgtacctgc aggcagaaag 240
tcttcctacg accgtctttt cccttagagg caccagaatc cctgtaacca ttcatccagg 300
tgttgagaag atatgtagca gccgagcacc catcttttga caccgtcctc tgaaatcagc 360
tttggagatg ctttcactct gtccgtcttc tgcagcagcc aggcagagtg ccgactcctt 420
cacagccgtg aggaactctt caggctccag aagctcttaa acctgatcta caatggaaaa 480
aattcttttt tatctgtttc tcattggcat agcagtgaaa gctcagatct gtccaaagcg 540
ttgtgtctgt cagattttgt ctcctaatct tgcaaccctt tgtgccaaga aagggctttt 600
atttgttcca ccaaacattg acagaagaac tgtggaactg cggttggcag acaattttgt 660
tacaaatatt aaaaggaaag attttgccaa tatgaccagc ttggtggacc tgactctatc 720
caggaataca ataagtttta ttacacctca tgctttcgct gacctacgaa atttgagggc 780
tttgcatttg aatagcaaca gattgactaa aattacaaat gatatgttca gtggtctttc 840
caatcttcat catttgatac tgaacaacaa tcagctgact ttaatttcct ctacagcgtt 900
tgatgatgtc ttcgcccttg aggagctgga tctgtcctat aataatctag aaaccattcc 960
ttgggatgct gttgagaaga tggttagctt gcataccctt agtttggatc acaatatgat 1020
tgataacatt cctaagggga ccttctccca tttgcacaag atgactcggt tagatgtgac 1080
atcaaataaa ttgcagaagc taccacctga ccctctcttt cagcgagctc aggtactagc 1140
aacctcagga atcataagcc catctacttt tgcattaagt tttggtggaa acccccttgca 1200
ttgcaattgt gaattgttgt ggttgaggcg tctgtccaga gaagatgact tagagacctg 1260
tgcttctcct ccacttttaa ctggccgcta cttttggtca attcctgaag aagagttttt 1320
gtgtgagcct cctctcatta ctcgtcatac acatgagatg agagtcctgg agggacaaag 1380
ggcaacactg aggtgcaaag ccaggggaga ccctgagcct gcaattcact ggatttctcc 1440
tgaagggaag cttatttcaa atgcaacaag atctctggtg tatgataacg gaacacttga 1500
cattcttatc acaactgtaa aggatacagg tgcttttacc tgcattgctt ccaatcctgc 1560
tggggaagca acacaaatag tggatcttca tataattaag ctccctcact tactaaatag 1620
tacaaaccat atccatgagc ctgatcctgg ttcttcagat atctcaactt ctaccaagtc 1680
aggttctaat acaagcagta gtaatggtga tactaaattg agtcaagata aaattgtggt 1740
ggcagaagct acatcatcaa cggcactact taaatttaat tttcaaagaa atatccctgg 1800
aatacgtatg tttcaaatcc agtacaatgg tacttatgat gacacccttg tttacagaat 1860
gatacctcct acgagcaaaa cttttctggt caataatctg gctgctggaa ctatgtatga 1920
cttgtgtgtc ttggccatat atgatgatgg catcacttcc ctcactgcca caagagtcgt 1980
gggttgcatc cagtttacta cggaacagga ttatgtgcgt tgccatttca tgcagtccca 2040
gtttttggga ggcaccatga ttattattat tggtggaatc attgtagcat ctgtgctggt 2100
attcatcatt attctgatga tccggtataa ggtttgcaac aataatgggc aacacaaggt 2160
caccaaggtt agcaatgttt attcccaaac taacggggct caaatacaag gctgtagtgt 2220
aacgctgccc cagtccgtgt ccaaacaagc tgtgggacac gaagagaatg cccagtgttg 2280
taaagctacc agtgacaatg tgattcaatc ttcagaaact tgttcgagtc aggactcctc 2340
taccactacc tctgctttgc ctccttcctg gacttcaagc acttctgtgt cccaaaagca 2400
gaaaagaaag actggcacaa agccaagtac agaaccacag aatgaagccg tcacaaatgt 2460
tgaatcccaa aacactaaca ggaacaactc aactgccttg cagttagcta gccgtcctcc 2520
cgattctgtc acagaggggc ccacgtctaa aagagcacat ataaagccaa atgctttgct 2580
gactaatgtt gaccagattg tccaggaaac acagaggctg gagttaatct gaagagcacc 2640
acttctcctc tctctcctga aaaaatttgc cactgatatt tttactggat aaaattcaaa 2700
aatgtttcaa ttcacaaagg ctaattgttg aactggtgtc gtagaagaaa ttgtctacag 2760
gagccaaggt gaaagtctct gatgacggcg gaactggctc cattagacca tggttcatcc 2820
tcttttaaaa ccaaattttt ttttcttctg gcctacaagt attttttttt taaaaaagaa 2880
aaaaagccta cattggcatc aagttctgta tcaatccatc ttacattgcc atccatgatt 2940
taacagactg tagaatcttg aataatctat atcactttaa caaataaatg ttttactatg 3000
acaaaaaaaa aa 3012

<210> 125
<211> 1600
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 7475633CB1

<400> 125
cgcgcgagga gcagccgctg ccgccgctcg gggtcgcctg gagccccaga ttcccgagcg 60
ctcggctcgc atggcagccg cttcggcgcc cggccccgcg gccagctagg ggcggccccg 120
cgctccctca cggcccctcg gcggcgcccg tcggatccgg cctctctctg cgccccgggg 180
cgcgccacct ccccgccgga ggtgtccacg cgtccggccg tccatccgtc cgtccctcct 240
ggggccggcg ctgaccatgc ccagcggctg ccgctgcctg catctcgtgt gcctgttgtg 300
cattctgggg gctcccggtc agcctgtccg agccgatgac tgcagctccc actgtgacct 360
ggcccacggc tgctgtgcac ctgacggctc ctgcaggtgt gacccgggct gggaggggct 420
gcactgtgag cgctgtgtga ggatgcctgg ctgccagcac ggtacctgcc accagccatg 480
gcagtgcatc tgccacagtg gctgggcagg caagttctgt gacaaagatg aacatatctg 540
taccacgcag tccccctgcc agaatggagg ccagtgcatg tatgacgggg gcggtgagta 600
ccattgtgtg tgcttaccag gcttccatgg gcgtgactgc gagcgcaagg ctggaccctg 660
tgaacaggca ggctccccat gccgcaatgg cgggcagtgc caggacgacc agggctttgc 720
tctcaacttc acgtgccgct gcttggtggg ctttgtgggt gcccgctgtg aggtaaatgt 780
ggatgactgc ctgatgcggc cttgtgctaa cggtgccacc tgccttgacg gcataaaccg 840
cttctcctgc ctctgtcctg agggctttgc tggacgcttc tgcaccatca acctggatga 900
ctgtgccagc cgcccatgcc agagaggggc ccgctgtcgg gaccgtgtcc atgacttcga 960
ctgcctctgc cccagtggct atggtggcaa gacttgtgag cttgtcttac ctgtcccaga 1020
ccccccaacc acagtggaca cccctctagg gcccacctca gctgtagtgg tacctgccac 1080
ggggccagcc ccccacagcg caggggctgg tctgctgcgg atctcagtga aggaggtggt 1140
gcggaggcaa gaggctgggc taggtgagcc tagcttggtg gccctggtgg tgtttggggc 1200
cctcactgct gccctggttc tggctactgt gttgctgacc ctgagggcct ggcgccgggg 1260
tgtctgcccc cctggaccct gttgctaccc tgccccacac tatgctccag cgtgccagga 1320
ccaggagtgt caggttagca tgctgccagc agggctcccc ctgccacgtg acttgccccc 1380
tgagcctgga aagaccacag cactgtgatg gaggtggggg ctttctggcc cccttcctca 1440
cctcttccac ccctcagact ggagtggtcc gttctcacca cccttcagct tgggtacaca 1500
cacagaggag acctcagcct cacaccagaa atattatttt tttaatacac agaatgtaag 1560
atggaatttt atcaaataaa actatgaaaa tcccaaaaaa 1600

<210> 126
<211> 1001
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1431268CB1

<400> 126
caatttttgt atctttagta gagacggggt ttcaccatgt tggccaggct ggtcttgaac 60
tcctgacctc aagtgatctg cccgccccgg cctcccaaag tgctgggatt acaggcgtga 120
gccactgtgc cccgccccga actcaggtct ttctgaccca ggagcagcac ctgcttcagc 180
cactgtcttt gggtccctgt ttggctgagt cacatctctc cctccatgtc taggctggag 240
tcctcagaag ctgcgtgcag ggctgtcccc tcagcctggc atactttcct cctgtcaccc 300
ctttgtctcc tccttattca agtctgggcc cacgggcctt ctctgcaggt cgtaactaaa 360
gtcgcacctc ctgccctaac ctccagcatg tctgactctt tggtattcac caagcacttc 420
tcactttgca aagtcattga ttctgcaaat gttcatcgag gatgtactac gtgccaggct 480
ctggttaagg cacgggatgt agaaacaagt tgctgtcgtt tttcagctca tgctctggct 540
ggagaggcgg tcagtcagca gaataagcaa agaggcggag aggctgcgtc ctgcctcctc 600
agatgagcac tgtgctgtcc agcacggtag ccgctggcca caggtgagat tttcacttaa 660
gttgaaatag gccgggcgca gtggctcatc atgcctgtaa tcccagcact ttgggaggcc 720
gaggcaggcg gatcacgagg tcaggagatc gggaccatcc tggccaacac agtgaaactc 780
cgtctctact gcccaggtgc tggtgcagag gaggcattcc atcaattgaa ccttaaggaa 840
ctctggaggc aggggatatg aatgtttact aagtgctgac atcacattaa gtacagaaat 900
cgtctggtgt cactacttca gtgtcaccgc tgcaacatcc agccgagggt gtcatgtcgt 960
ttgcagagca ggaaactcag cctcagaatg gttgccttgt c 1001

<210> 127
<211> 1424
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2414185CB1

<400> 127
tcatggggtg gaataaacac agaaaagact gaaaaagccg ggccaatggc tcaagcccgt 60
aattcccagc acttgggagg ccaaggtggg cggatcacct gaggtcagga gttcaagacc 120
agcctggcca agatggtgaa atcccatctc ttactaaaaa tacagaaatt agccaggcat 180
ggtggtgggt gcctgcaatt tcagctactc aggaggctga ggcaggagaa ccgctggaac 240
ctgggaggcg gaggttgcag tgagccgaga tcccgccact gcgctccagc ctgagccaac 300
aacagcgaga ctccgtctaa aaaaaaaaaa agaagactga aaaaaaatat gttggagacg 360
actttcttac tgtggaaggg ctgagattcc atgatgggaa agcttagccc caccttttatt 420
ctgggcattt gttgggtacc tgctggtttg ggataccagc agggaaagaa aacatggccc 480
ttgtcctcca gcccttacaa cctgcaagac aagatgtacg ccttggagaa agctggagat 540
ccttccaagg caaggagcat gggccctcac aaatctccag aaacccagag agggcaacct 600
atggagatgt ctggcttaaa gggccaggtg acatccacag ccttacacac tctgcatttt 660
cctagaaggc ctcccctctgg ctgtcagaca gatcaggctg gtgaccacga gccaggggga 720
cgctttctgg cccaaccaca gagacttagg gaactttcat tgatgatcag tcccctgcaa 780
cttcttccct ttggtagccg ctgaaagtag gtcatgaaag ggcctagacc tcaggatgtc 840
tggggagaac ctccttgagt tgccagaaga ttcttgcatc cagggctttg tggctaaggc 900
tcatgcttac tggcagtgct ggcctgatgg gtagagaacc gcctggacca cctacgtggg 960
atatttatgg gatgctgaga cgtttctcca aaataccttc tggaatggcc tgtatcagta 1020
gctgaagtca taggatcatt cctcccagct aaagaaatgc ctgctattaa atttttgtgg 1080
ataagaggaa gagacatagc ttgtggacgt gcagttgggt gaagtcatga acatactgac 1140
taatgatacc cagggtgaac tgctgactaa tttgtgtaaa cagagtggat ggtctctgat 1200
ggtgagccct agggctctgt tccggcctcc ccatcatgta tcaggaatgt tcatggaaat 1260
taaaaggtat actgataaaa tctatgggag atacaaaact aggagagatt gctggagtaa 1320
taaaattatg gaatgaactt tcagaaaaac tgctgtaggt tgaaataatg ggcccaagct 1380
gtaacatatt agggatgaat agaaggtcct atatcaaagt tcat 1424

<210> 128
<211> 1282
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5266594CB1

<220>
<221> unsure
<222> 9-10, 23
<223> a, t, c, g, or other

<400> 128
aacaaaaann cccaaaacca gtnaaaagcg tgcgctgacg gtttgccatt taaggattac 60
agcaaatgat tctatttgta atttctaccc ttcccattgc ctccccccca aaaaaaactg 120
tacatgagtt tacaaacata ttaacatata aataatgaga accgtcctgg tgggagcctc 180
ctccgttgtc tctgctggag atgaacactg aggggcgctg taaccacaca gactgcctgt 240
gacatcggga gtctcacggc agctgtcctg ggcccgcagc tggctttttt ggcacctcca 300
ggttcaacca ccagtctgtc tctgctgtgc ccagggtaga gcccgggggc tgtgagtatg 360
tgtggctccc ctgcccgtca tcgctctggc tcaagctcat gctggaaggg acgcttcctc 420
tcccgacagt gcttcttgtg ggcaggccag tccttctgct ggcactggga gccgcagtac 480
cgggccacct ggcagcgccc acagatgttg aactcccgga gctgcttctc aatcactgtg 540
cagggagggt agtggcactc atagtaggtg caagagttct cctcctcttc cactacatcc 600
ccgttggcat tataataccg ggtcacattt aggacaggaa actgttcttc tatagggtct 660
gtaggcagct gctgaattgc aagccaatac aagctttgct ccctctgctt gctggaggtc 720
tcttcagcct tgggccgcca gccccacggc accagttgca ggttgtccag gcgattgtcc 780
acggtcacag cgttgaggtg caccacctgg aagcccgggg ccacgccccc ccggtgccgc 840
tcccacagca gctcatgaag gagtctccca gagccccttc ctcggttctt gtcaaaggca 900
taggcaaata tcttagcacc atttccatct gcatccactt ccattcgggc ctcaaaggag 960
tagctctcca ccagcgggat gtcctgctca tcgatcagcg tgtatttggt cttcccggcc 1020
acccggccga gccgcacgat acccaatttg aagtcggtca tggccgggcc tgcgcttcgc 1080

cggcagcccg ctccctcggg aggcgccgac ggggccgggg cgagccgcgg cgcgccggta 1140
cagacgggac cgagccgggt cgggtacacc agcgggctcc gggcggacga gctggccggc 1200
tcggcctccc gccgcctgcg cccgccgctt cagaactgcg ccgtcgagat gattaaagcc 1260
atgccctata tatattacgc gc 1282

<210> 129
<211> 642
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7610617CB1

<400> 129
cggctacagc tgcaagtcaa agaggtcaga ggccttctga tgttggcgct ggaagcctct 60
cctttcttgc cccacaggtt atgggggctt ctccactccc cgctcttact tgcctgcttc 120
tccagggatc cctagaagag cagtggttag caagcatatg ctatagccat caaatacgag 180
tctacagaaa tgtaggcttc cccttggggc tgagttaggg cctctgcttc tgaaatgagc 240
aacctctggt tgctggtggg ggctagggct tgctcccttt ctctcctcac atactctttc 300
ctgggtgacc tcatcccctc ccattgcctt aaacatcttc caggtactgg tgtcatccac 360
ttgtgtagtt ccagttcaga aatcccgtca gctccattca ttcatttgtt cattcattca 420
gcaaacatat gtggcatctc tgtcccaggt actgctctac aacctgggtg tacaataggg 480
acccaaacag acactccctt ccccatgcac tccctgctca cagacacacc tgcgtggcag 540
tgcctgggtg tcttcacagc tccctgatgc cgagctgtga tcccaccctt ctgcctccct 600
ggtccccctc agcgtgctgc caccattcac cttgaggctg ag 642

<210> 130
<211> 1326
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1902436CB1

<400> 130
ctcacctcag cctcctaagt agctgggact acaggtgcct gacaccatgc ctggctgatt 60
tctgtttgtt tgtttgtttg tttttgtaga gatggggtct tactttgttg cccaggctgg 120
tctcagaaat gacttttata tatactttta ttttaagttt ttttcttcaa ctatgttgtt 180
cttttatgaa gttaatcttg ctcatcagca acacaaatgc tgtatctttt atactacaca 240
ggccctgtac cctttgctct gatttctact cccatatttg tatgctgtta actgtctctg 300
tgaatttcct gtcattttgg aataatttttc aaactattct tacttgggct gatctgtttt 360
caatgctttt ggcatacgaa tatagattta caagactctt ctcagtgcta ccccacactt 420
ctgtcatgtt atgtttttaa aaattctctg attgagctat aactgcttaa agatgtgaaa 480
tgagtatttt caatagattt agtttttttca gggggtaatg acagacttgt taaaaaaaaa 540
aggttataaa tacaaaaagt ataaagaagg aaaaacctat gcaaaatctg agcagagtta 600
actttctggt gaatgtcatt ccccgtctct agttatgtgt gtataattat gcataataga 660
gattctattg tatccaagga ttttttaaac ctttttgagg tgagatctgg ctgtgttgcc 720
caggctggag tgcagtggct gttcataggt aagatgatgg tgcactatag ccaaggactc 780
ctggactcag gtgatactcc tacctcagcc tcccaagtaa ctgggactgc tggcacacac 840
cactgcacct ggcttaagtg ttttttccaa gaagtcaaag gctaatgatt ctaaagaaag 900
ttagtgtaca gacccttata catgagtatt atgtgcatct tctttagtgg cagaattctg 960
ttttaagtat tataattgga aaaacaactg aagagacatt taaggaacac cactatttca 1020
gcagacagtt gatagaagat acctttttaaa agaagtggtg gaatgtcaca catttgaaac 1080
ttaaaaaaag ataaacccaa aacgtcagta gaagggatgc cagagggaat aatcttgatg 1140
gcaagcagat agatttgata agccaatggg atttcccact tctctttgct atttttctat 1200
tttgcagccc tacttgattt aggcaagtat catgagtgga tgctaaaagt acttagtaaa 1260
agtttgatga ggaatagaat gtttattcaa tcccaaagta tactcatatt atcaccaaaa 1320
aaaaaa 1326

<210> 131
<211> 1486

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2310369CB1

<400> 131
cccaaacaga gcgttttctg gtcgccgaac tcctcccggc tccacccgca cctacagcct 60
ccaaccaaac cccaccaagt ctcgaccctc tcccccaggc ctcgaccctc ccgctcgccg 120
ctgcgttggg accgcgcacc cggtaactgc gcacggctgg cctcccggag tcctgcgccg 180
tcgctcctcc tggacccggg ttcccaggac aacgcctccc ggaacgcagg gagcaggccg 240
aggccgccgc gtgggcctgc agcgcctcgc cgctctctct tccccaggcc cgcactccca 300
cttgggctcc agggccccag ggctgacgtt ccccccagct tagaccctga gtcgttttcc 360
cccgtttccc ggctgaatta ggttcttctt ctccacaggt gtgtgcagtg gcctcaggga 420
tccggaaagt ctaggactga acttctccta acatccagta atggggacct ggaacctggg 480
cgtactagag tgccgcgcgt agggctccag gtcgctggct tctgcgcttt cttcctctcc 540
aaagttgagt atctcctatc tgtgtcctca tacatactgc cgcctgaggt gccatggccc 600
ccaagccggg ggccgagtgg agcacagccc tgtcccatct ggtgctggga gtggtgtctc 660
tgcacgcagc cgtgagcaca gccgaggcaa gtcgaggggc tgctgctggc ttcctgctcc 720
aggtcttggc tgccaccacc acgctggccc cagggctgag cacacatgaa gactgccttg 780
ctggagcctg ggtggccacc gtcatcggcc ttccccttct ggccttcgat ttccactggg 840
tgaatgggga ccgctcctct gccaacctgc tcctgggagg aggcatggtg ctggcagtgg 900
ctggcggcca cctcggccct gagggccgct ctgtggctgg tcaggcaatg ctgttggtgg 960
tcgcagtgac catcctcatt gtagctgtct tcacggccaa cacttatggg atgtgggggg 1020
gggcgatgct gggtgtggca ggcctcctga gccggctgga ggaggacagg ctgctgctgc 1080
taccgaagga ggatgtctgt cgctgggcct tggctgtagg cagctgggct tactgccggg 1140
ccctgcatac acagcgcctc cagtgggagt gacagttgga tacagccagg cagggtttct 1200
gccctgccga acactttccc tcccacctgc ctgctcctgg cgccttctcc ctaggggtag 1260
actcttctgc ctactgaagt gggtttgctg cacattgact ggtcaggggc agagtctggg 1320
tgctgtcctt tggccacgtg tggggacttg tctagaccag aatgaaaggg acagggtccc 1380
agacacgttt gggggtcctg attctgggct ggacacggtt gtggatccag agaagaggcc 1440
tagtctccaa taaatcttag gaattttgca ggaaaaaaaa aaaaaa 1486

<210> 132
<211> 1523
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6180576CB1

<400> 132
gcggcggtgg cggtgcagga ggccgggcag gggtgcggag ggaccgacgg acgcacgggc 60
gggcggccgg gagccatgga gcgcggccct ggggcccggg ggcgcgggcc ggggtgggct 120
tcccacggca cgacatggag acctgtggtt gcgaggctcc ctggggctcg gcttggaccg 180
cgatggggct gggccctggc ctcctaacgg ggctgctgtc tggggcggta gctgggggag 240
cgctctcccc cctgcccgcg actcggagca cccccacccc tcccctgccg ggccaggccg 300
ggcggcgttg ttggcggggg ccccggtgga ggcccggcct gggcggcgcc cgccatgaat 360
gggctgtcgc tgagtgagct ctgctgcctc ttctgctgcc cgccctgccc cggccgcatc 420
gctgccaagc tcgccttcct gccgccggag gccacctact ccctggtgcc tgagcccgag 480
ccggggcctg gtggggccgg ggccgccccc ttggggaccc tgagagcctc ctcgggcgca 540
cccgggcgct ggaagctgca cctgacggag cgtgccgact tccagtacag ccagcgcgag 600
ctggacacca tcgaggtctt ccccaccaag agcgcccgcg gcaaccgcgt ctcctgcatg 660
tatgttcgct gcgtgcctgg tgccaggtac acggtcctct tctcgcacgg caatgccgtg 720
gacctgggcc agatgagcag cttctacatt ggcctgggct cccgcctcca ctgcaacatc 780
ttctcctacg actactccgg ctacggtgcc agctcgggca ggccttccga gaggaacctc 840
tatgccgaca tcgacgccgc ctggcaggcc ctgcgcacca ggtacggcat cagcccggac 900
agcatcatcc tgtacgggca gagcatcggc acggtgccca ccgtggacct ggcctcgcgc 960
tacgagtgtg ccgcggtggt gctgcactcg ccgctcacct cgggcatgcg cgtcgccttc 1020
cccgacacca agaagaccta ctgcttcgac gccttcccta acatcgagaa ggtgtccaag 1080
atcacgtctc ccgtgctcat catccacggc acggaggacg aggtgatcga cttctcgcac 1140

gggctggcgc tctacgagcg ctgccccaag gcggtggagc cgctgtgggt ggagggcgcc 1200
gggcacaacg acatcgagct ctacagccag tacctggagc gcctgcgtcg cttcatctcc 1260
caggagctgc ccagccagcg cgcctagcgg cggccccaac cggccggacc tcagcaataa 1320
ggcggccccc ggacctcacc ccgcgccggc ccccacccag gggctgcatg tggacccccc 1380
gggcggccca ggggaccccg ccccgaccca ggggctgtgg acgatgtaca ggcaacagag 1440
ctacgcactc ctttcctttt ggaagcaaga agaaaatacg tgaaaacgga aattaaagat 1500
ttaaaatttt aaaaaaaaaa aaa 1523

<210> 133
<211> 848
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2274523CB1

<400> 133
tagttctaga actcctgacc tcctgatctg cccacctcag cctcccaaag tgttgggatt 60
acaggggtga gccaccacgc ccagcccagc agtggcttct gtgtgggaag aaggtcaggg 120
ctgaaggacc ctgggaggat gccccaccca acatgtcctc tctcatcgtg gggtaattgc 180
ccatgtttgc ccagccattc tcacctatca gggcttccaa aagaatggcc aaagtaagtt 240
ctaacaactt tgcttctttg ccaaggcaag caccaatgct acttttctgc cctctctgga 300
tgcccgtgac ttcggtccca caggaagcca agctcctgag gcagctgaaa ttcagtcagg 360
ggactggggt ctgtgtactg atctacaccc cacttcatac ctacttcttc aaactgtccc 420
caacactggg aacacctgtc cttgaatatc ctcagcaatt taaggggaaa aaaaggctga 480
agcagaaaga ctttttcctt ccaaaattgt gcctgcttgc ttgggggcct aggcatgcag 540
acttgaaaat aaatcaagct tgggtggggc atggtggctc ccgcctgtga tcccagcact 600
ttggaaggcc gaggcaggcc aggcagggtg aggaggatca caaggtcagg agatcgagac 660
catcctggct aacacggtga aaccctgtct ctactaaaaa tacaaaaaca aaattagcca 720
gggtggtggc gggtgcctgt agtcccagct gctctggagg ctgaggcagg agaatggcat 780
gaacccagga ggcaaagcta gcagtgagcg gagatcgcac cactgcactc aagcctgaca 840
taattgcg 848

<210> 134
<211> 2758
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1801820CB1

<400> 134
cggcggggcc ggcgcctgcc tggagggatg gggctgccgg gcgcgtaggg gccatgccgc 60
ccgggacccg ggcctgccgc gttccgcgcc ccggccgccg cgccccacgt ccgcgccggg 120
atggtgaacc tggcggccat ggtgtggcgc cggcttctgc ggaagaggtg ggtgctcgcc 180
ctggtcttcg ggctgtcgct cgtctacttc ctcagcagca ccttcaagca ggaggagagg 240
gcagtgagag ataggaatct cctccaggtt catgaccata atcagcccat cccgtggaaa 300
gtgcagttta acttgggcaa tagcagtcgt ccgagcaatc agtgccgcaa ctccattcaa 360
gggaagcacc tcatcacgga tgaactcggc tacgtttgcg agaggaagga tttgctggta 420
aatggctgct gtaatgtcaa cgtccctagc acgaagcagt actgctgtga tggctgctgg 480
cccaacggct gctgcagcgc ctatgagtac tgtgtctcct gctgcctgca gcccaacaag 540
caacttctcc tggagctctt cctcaaccgg gcagccgtgg cattccagaa cctcttcatg 600
gcagtcgaag atcactttga gttgtgcctg gccaaatgca ggacctcatc tcagagcgtg 660
cagcatgaga acacctaccg ggaccccata gcaaagtatt gctatggaga aagcccgccc 720
gagctcttcc ccgcttgacg ggtgcagcgg acttgctcca gcctgggtga ggaggccccg 780
ctgaagaact cgcctcctgg gacccagctt cagccatcgg gccaggctgc aggaagaaga 840
caaaggcagc gtgaggaaac cttggctttg acccccttctc gtgttgtcat ctttggcttc 900
gctcaccacc cgggcttacc agatggaact cttctgtaaa gcagcttggc ccctccagcc 960
agtcccattc gggaaagatg aaaccggagg ccgggctcac ggtggtggtg gagttcttgg 1020
atgactcagc cctgggacct gcacagggac ctgtgacttg tgttcatcgg gggccggtgt 1080
cacttccagt tttgatccag gctctttcac tgtaaaatta tttattggat tcctttggag 1140

taatgggaac attctaatgt tttatgtagg aaaatgcctt gccattctag ttgaatatgt 1200
tcaaggaaat tatttttgtt gttgttctgt gttctcgagt ttcaggagtt aaatcattct 1260
tccacccaga tacaacattt tctcttttag gacgtgaata ttctctctag gcagttattt 1320
ttgtttgtat ttcgacagta tcaagcatag gccctgaacg tgacctgtta gccatatcct 1380
gacgtgtaaa attatctaaa aactcagaca ctcttccatt ctaatctaac tgcacgattt 1440
ctaacagtgg gcaccatgtg cctgccctca ggttattttc cagtggtagt cgaatgtgct 1500
catataccct atggagagca ctgttttagc agaaatctaa tttctcttcc tggaggaatt 1560
tgttctcatt tcttttgcca cttaaaatta actgtgggct actcagccag ggtacagtgg 1620
gagcctcagg aaggtcagag gcaacctcct cccctgttct atcaatagaa acccaacgtt 1680
gaggcaattc ctaaacagac gcacctcgta gcttgctgta tgtgtttatt ctttattgct 1740
ttcagctttg gggctgtaac aggtacaaat atttggtttc cctatgattt atagagaaga 1800
agaagaaacc cagctttcta tcagagcact gcaagagaag agtcttacac ctgccctcag 1860
tgggagatga gaatggtcat tatgacttag agaatgctac acgtgtaggt tgctggtgtg 1920
tcctgaatcc acaggcataa agcactcccc atttttccta ctgtaatgca gattctccgg 1980
ctcaaggtct agaatatttg atcctaagat caagacatca tgcccttcga atagtactgc 2040
tctttgtttt caggagtcac gtgaacacac aactctccta tattcctcac gaacctcagg 2100
attgagcaag gtctttgtaa ttttttggt tcactttatt gacctgggag caaggtgcta 2160
attctgtggt cagtattcaa tgtttttttc agtggagctt tttctttggg ccatatttgc 2220
cttctaatac attcctgcaa tatgtagtgg tgatttccct tagcttcctc ctactacctc 2280
ttatactcat ctccccaaat tatttgcctc ccttaaataa gttttcctag aaggtaagct 2340
ggtcaggcaa tttgaaaaat attagatccc aagaaatcta ttccgtttgc attggacttc 2400
tcggattcca tgtgtttgca gcaggactac atcgaactct gatgtgccgg attgtggcat 2460
gtctgcatgt ctcatccatc tattgttttt ggtaactcag tttggaattt cagtgtctgt 2520
cttccctggg ttgacattgg aatcagcctc tcctttgagc ttattttaac tcttgagcaa 2580
cataacatag atttaatgtg aacagtttat accaaagggc agcctgtgcc tgtttatgga 2640
tcctctctgc ctttgtactt gaagagcgca ttttacattt ccagtccttt cacagacagg 2700
agctccaacc ttacgatgga gaattaaact tgcttgtatt tccactttgt ggatgagg 2758

<210> 135
<211> 1650
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3010920CB1.comp

<400> 135
gtcagcatta cttatgtttc atcctgtgtt cgtgtgcttt ctcgctgttg ctacattcaa 60
catccaaatt gtaatttata tttataactg cctctcatct aattatattc tttctcgttt 120
ttaaaacaca gcaaaatcaa ctttagctag ccattatact aaaacatact acattttctt 180
ttaaacttct tttagtttgt ttcctctaga tggtaaactc tactcaggca gggattatgt 240
ttgattattg atgtgaatgc aacttttag aacatagctg cgtactcagt ggactcagta 300
catactgacc agttgttgaa aaatctgatg cagcctccct aagtgtttag catagccaga 360
gtagtttacg tattcgttaa atctgactca gtaattagca gacctagtgt ttaagtactg 420
gggaaatacc tgcttgccтc tcacagatgt caagaaagta cctggtagaa gagtcttacg 480
acacagccta acttgtgtag gacagaagat accatctgga ctttgtagaa cccagaaata 540
caagtctaaa aatatcttga taacaaattc tttgaaatat ttttgaaggt cctgagcact 600
tcacttttt ggctatatta ttaggaattg tttgatctct gcaatttcaa caatgattac 660
agtaatcact gtcattatca agcatttact atgggccaga ccctttccta ggtactttac 720
atacaatata ttgttctgtt aaaaacttaa caaggtttgt attattattt tacaaatgaa 780
gaaactaagg ccgagggagg ttaactaact tgcccaggat cacacgtatg gtaaaggcag 840
aagtaaatac ttgagatatg gtttgttgag ttctaaaact taggtgaatt ttacctccta 900
atacgtctca aggatgtggg tactggagga agaactgtac tgaaatccaa atatttggtg 960
ctagctgttt tgtataatac tagtaaatat tcatattgcc cttatggtat tgcttatgat 1020
catacaataa ttggttcttg cttatataga aatctacagt tttctgtagt gaatatatgt 1080
gttacagcaa ccacaactat ttatctgaaa tatcagtaaa tatttatcat aggaatgaac 1140
atatattaag tcttgacaga gctctcgagg cataatgact tggttgaaag tagcagattt 1200
cagtaagaac catgccaaaa atattaagct ataagttatt agtgtaccta gttttaatgg 1260
caaacatgat gattttctta gtgccatatg atgaaataaa taacttacta gaagtttgta 1320
gacaggtttt tttttccccc tctggagaca gagtcatgct ttgttgccca ggctggagtg 1380
caatggcata attttggctc actgcaacct ctgcctccca ggttcaagca attctcctgc 1440
ctcagcctcc cacgtacctg ggattaccat gcctggctaa tttttgtatt tttagtagag 1500

atgaggtttc accatgtttg cctggctggt ctcgaactcc tgacctcagg tgatccccc 1560
gcctgggcct cccaaagtgc tgggattaca ggccacctca cctaggctag atagttttat 1620
tcacttggct gtttcaccaa aaaaaaaaag 1650

<210> 136
<211> 1061
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3409459CB1.comp

<400> 136
aaaataaagg ttcaagctgc caaattttct tccagggtct gtttgctctc cctccagaga 60
tgtgctgagt ggccgtggct tcttcccacc caccgtgtcc tgctcttccc tgttcttggc 120
ctaagtcttg cacaggagcc tgtcgaggcc cttgggacac ttccagaagc caggtccatt 180
ttctccatga acaggtgggc cagggttgag tctcccccta gttggactgg gagcacatgt 240
ggtaagatgc caggggaagc tccgtgttta tcctccacat gcctggaagg atgcagtgga 300
agcctttgcc tggcgtggtt tctttctggc ttgaagtttc tgccaagctt tagggccctg 360
ctggcctggg aagcagcgtt ggcaggcctg ccttgcggat ggcacgctgc agcactggca 420
gggggcagtg ccaggcggga ggggcagggg aggatgccct tcctggccaa atcccttggt 480
taagtaagta ccagcatctg aggtggcccg ggtgcagggg aaccactatt gctacctctc 540
cagcagcagc cagtggtgag tctgacatgt cagctgcatt cagaccaaac ccgccccccg 600
caaataccca gagactgagg caggaagctc tgtagaaatt cattttattc ttattcagac 660
tattttcaaa agaagcagtg gtgtgctgtt tttctaaaaa tatgccttta tagattttta 720
tatatgtata ttataaaatc catacatgta tttacatgat tgctacatac aaaattacag 780
cactgtggta tgtacacatc tacaggtaca ttcttgccgc gcatccctgc tgtgctttcc 840
ccacgtgagg gagggaggga gactgaatcg gttgtgagca gctgagggct ggccgggccg 900
cggagccttt gatttggggc ctgggtggag gaggaggggc ggctgcttca catcactctc 960
cagctctgac actcggcctt cccactcccc ttctcggttc tcaaatcgtc ttctcaattc 1020
atgtttctct tgctctaagt gctcgagctt gtctgtcagc t 1061

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 April 2002 (04.04.2002)

PCT

(10) International Publication Number
WO 02/026982 A3

(51) International Patent Classification[7]: C12N 15/12, C12Q 1/68, C07K 14/47, 16/18, A01K 67/027, A61K 38/17, 39/395, B01J 19/00, G01N 33/50, 33/53

(21) International Application Number: PCT/US01/30042

(22) International Filing Date:
25 September 2001 (25.09.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/236,869	29 September 2000 (29.09.2000)	US
60/239,812	11 October 2000 (11.10.2000)	US
60/240,108	12 October 2000 (12.10.2000)	US
60/241,282	17 October 2000 (17.10.2000)	US
60/242,218	20 October 2000 (20.10.2000)	US

(71) Applicant *(for all designated States except US)*: **INCYTE GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94087 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **NGUYEN, Danuiel, B.** [US/US]; 1403 Ridgewood Drive, San Jose, CA 95118 (US). **YAO, Monique, G.** [US/US]; 1189 Woodgate Drive, Carmel, IN 40633 (US). **XU, Yuming** [US/US]; 1739 Walnut Drive, Mountain View, CA 94040 (US). **TRIBOULEY, Catherine, M.** [FR/US]; 1121 Tennessee Street #5, San Francisco, CA 94107 (US). **SANJANWALA, Madhu, S.** [US/US]; 210 Sylvia Court, Los Altos, CA 94024 (US). **CHAWLA, Narinder, K.** [US/US]; 33 Union Square, #712, Union City, CA 945877 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **SAPPERSTEIN, Stephanie, K.** [US/US]; 3531 Highland Avenue, Redwood City, CA 94062 (US). **LAL, Preeti** [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). **THORNTON, Michael** [US/US]; 9 Medway Road, Woodside, CA 94062 (US). **GANDHI, Ameena, R.** [US/US]; 837 Roble Avenue #1, Menlo Park, CA 94025 (US). **RAMKUMAR, Jayalaxmi** [US/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). **ELLIOTT, Vicki, S.** [US/US]; 3770 Polten Place Way, San Jose, CA 95121 (US). **ARVIZU, Chandra** [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). **THANGAVELU, Kavitha** [IN/US]; 1950 Montecito Avenue #23, Mountain View, CA 94043 (US). **GIETZEN, Kimberly, J.** [US/US]; 691 Los Huecos Drive, San Jose, CA 95123 (US). **DING, Li** [CH/US]; 3353 Alma Street #146, Palo Alto, CA 94506 (US). **AU-YOUNG, Janice** [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US). **TRAN, Bao** [US/US]; 750 Salberg Avenue, Santa Clara, CA 95051 (US). **POLICKY, Jennifer, L.** [US/US]; 1511 Jarvis Court, San Jose, CA 95118 (US). **LEE, Sally** [US/US]; 825 East Evelyn #425, Sunnyvale, CA 94086 (US). **LU, Dyung, Aina, M.** [US/US]; 233 Coy Drive, San Jose, CA 95123 (US). **BURFORD, Neil** [GB/US]; 105 Wildwood Circle, Durham, CT 06422 (US). **WARREN, Bridget, A.** [US/US]; 10130 Parwood Drive #2, Cupertino, CA 95014 (US). **GURURAJAN, Rajagopal** [IN/US]; 5591 Dent Avenue, San Jose, CA 95118 (US). **DUGGAN, Brendan, M.** [AU/US]; 243 Buena Vista Avenue #306, Sunnyvale, CA 94086 (US). **HONCHELL, Cynthia, D.** [US/US]; 400 Laurel Street #203, San Carlos, CA 94070 (US). **HAFALIA, April, J., A.** [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US).

(74) Agents: **HAMLET-COX, Diana** et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
25 September 2003

[Continued on next page]

(54) Title: SECRETED HUMAN PROTEINS

(57) Abstract: The invention provides human secreted proteins (SECP) and polynucleotides which identify and encode SECP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of SECP.

WO 02/026982 A3

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/30042

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C12Q1/68 C07K14/47 C07K16/18 A01K67/027
A61K38/17 A61K39/395 B01J19/00 G01N33/50 G01N33/53

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C12Q C07K A01K B01J G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
BIOSIS, EMBL, SEQUENCE SEARCH

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	DATABASE EMBL 'Online! EBI; 23 May 2001 (2001-05-23) OTA, T. ET AL.: "Human membrane or secretory protein clone PSEC0241" Database accession no. AAB88452 XP002222315 abstract	1-20,23, 26-47, 56,123
P,X	-& EP 1 067 182 A (HELIX RES INST) 10 January 2001 (2001-01-10) *SEQ ID NO:272* the whole document	1-20,23, 26-47, 56,123

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 November 2002	08.05.03

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Hillenbrand, G

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	DATABASE EMBL 'Online! EBI; 23 May 2001 (2001-05-23) OTA, T. ET AL.: "Human membrane or secretory protein clone PSEC0241" Database accession no. AAF93879 XP002223091 abstract	1-20,23, 26-47, 56,123
P,X	-& EP 1 067 182 A (HELIX RES INST) 10 January 2001 (2001-01-10) *SEQ ID NO:271* the whole document ---	1-20,23, 26-47, 56,123
T	DATABASE EMBL 'Online! EBI; 6 November 2001 (2001-11-06) OTA, T. ET AL.: "Human polypeptide and full-length cDNA" Database accession no. AAM93427 XP002222316 abstract	
P,X	-& EP 1 130 094 A (HELIX RES INST) 5 September 2001 (2001-09-05) *SEQ ID NO.:3054* the whole document ---	1-20,23, 26-47, 56,123
T	DATABASE EMBL 'Online! EBI; 6 November 2001 (2001-11-06) OTA, T. ET AL.: "Human polypeptide and full-length cDNA" Database accession no. AAK94348 XP002223092 abstract	
P,X	-& EP 1 130 094 A (HELIX RES INST) 5 September 2001 (2001-09-05) *SEQ ID NO.:3053* the whole document ---	1-20,23, 26-47, 56,123
P,X	DATABASE EMBL 'Online! EBI; 8 February 2001 (2001-02-08) SHIMKETS, RA. AND LEACH, M.: "Human secreted Human ORFX ORF1664 polypeptide sequence and polynucleotide sequence" Database accession no. AAB41900 XP002222317 abstract	1-20,23, 26-47, 56,123
P,X	-& WO 00 58473 A (CURAGEN CORP ; LEACH MARTIN (US); SHIMKETS RICHARD A (US)) 5 October 2000 (2000-10-05) *SEQ ID NO.:3328* the whole document --- -/--	1-20,23, 26-47, 56,123

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	DATABASE EMBL 'Online! EBI; 8 February 2001 (2001-02-08) SHIMKETS, RA. AND LEACH, M.: "Human secreted Human ORFX ORF1664 polypeptide sequence and polynucleotide sequence" Database accession no. AAC76109 XP002223093 abstract	1-20,23, 26-47, 56,123
P,X	-& WO 00 58473 A (CURAGEN CORP ; LEACH MARTIN (US); SHIMKETS RICHARD A (US)) 5 October 2000 (2000-10-05) *SEQ ID NO.:3327* the whole document	1-20,23, 26-47, 56,123
A	WO 99 40100 A (HUMAN GENOME SCIENCES INC ;ROSEN CRAIG A (US); KYAW HLA (US); LAFL) 12 August 1999 (1999-08-12) the whole document	1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/30042

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

see FURTHER INFORMATION sheet PCT/ISA/210

2. [X] Claims Nos.: 21-22, 24-25,48-55
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-20, 23, 26-47 (partially), 56, 123

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claim 19 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Although claims 33 and 35 are directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Continuation of Box I.2

Claims Nos.: 21-22, 24-25,48-55

Present claims 21-22 and 24-25 relate to an extremely large number of possible agonists or antagonists and methods for use of this agonists or antagonists. Claims 48-55 are extremely broadly and imprecisely drafted product claims directed to an array comprising different nucleotide molecules. In fact, the claims contain so many options and variables that a lack of clarity and conciseness within the meaning of Article 6 PCT arises to such an extent as to render a meaningful search of the claims impossible.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: Claims 1-20, 23, 26-47 (partially), 56, 123

Invention 1:

An isolated polypetide of SEQ ID NO: 1 and polynucleotide SEQ ID NO: 68 encoding such a polypeptide (and fragments thereof), a promoter linked to such a polynucleotide, a cell or transgenic organism transformed with such a polynucleotide, a method for producing such a polynucleotide, an antibody (and method of its preparation) that specifically binds to such a polypeptide, a method for detecting a target polynucleotide in a sample, a method of treating a disease comprising administring to a patient such a polypeptide, a screening method for agonists or antagonists or compounds that bind to such a polypetide, a method for screening for a compound that modulates the activity of such a polypeptide,a method of screening a compound for effectiveness in altering expression, a method of assessing toxicity of a test compound, a diagnostic test comprising said antibody, a microarray comprising such a polynucleotide, and a method of generating a transcript image by using such a microarray.

Inventions 2-67

An isolated polypetide of SEQ ID NO: 2-67 and polynucleotide SEQ ID NO: 69-134 encoding such a polypeptide (and fragments thereof), a promoter linked to such a polynucleotide, a cell or transgenic organism transformed with such a polynucleotide, a method for producing such a polynucleotide, an antibody (and method of its preparation) that specifically binds to such a polypeptide, a method for detecting a target polynucleotide in a sample, a method of treating a disease comprising administring to a patient such a polypeptide, a screening method for agonists or antagonists or compounds that bind to such a polypetide, a method for screening for a compound that modulates the activity of such a polypeptide, a method of screening a compound for effectiveness in altering expression, a method of assessing toxicity of a test compound, a diagnostic test comprising said antibody, a microarray comprising such a polynucleotide, and a method of generating a transcript image by using such a microarray.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 01/30042

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 1067182 A	10-01-2001	AU 5849300 A	30-01-2001
		AU 5849400 A	30-01-2001
		AU 5849500 A	30-01-2001
		AU 5850500 A	30-01-2001
		EP 1067182 A2	10-01-2001
		EP 1197554 A1	17-04-2002
		EP 1203816 A1	08-05-2002
		EP 1201754 A1	02-05-2002
		WO 0104312 A1	18-01-2001
		WO 0104299 A1	18-01-2001
		WO 0104300 A1	18-01-2001
		WO 0104301 A1	18-01-2001
		JP 2002017376 A	22-01-2002
EP 1130094 A	05-09-2001	EP 1130094 A2	05-09-2001
		JP 2002017375 A	22-01-2002
WO 0058473 A	05-10-2000	AU 3774500 A	16-10-2000
		EP 1165784 A2	02-01-2002
		WO 0058473 A2	05-10-2000
WO 9940100 A	12-08-1999	AU 2577799 A	23-08-1999
		CA 2320625 A1	12-08-1999
		EP 1053245 A1	22-11-2000
		JP 2002502589 T	29-01-2002
		WO 9940100 A1	12-08-1999